Statistical Physics

A. ISIHARA

STATE UNIVERSITY OF NEW YORK, BUFFALO

ACADEMIC PRESS *New York and London*

COPYRIGHT © 1971, BY ACADEMIC PRESS, INC.
ALL RIGHTS RESERVED
NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM,
BY PHOTOSTAT, MICROFILM, RETRIEVAL SYSTEM, OR ANY
OTHER MEANS, WITHOUT WRITTEN PERMISSION FROM
THE PUBLISHERS.

ACADEMIC PRESS, INC.
111 Fifth Avenue, New York, New York 10003

United Kingdom Edition published by
ACADEMIC PRESS, INC. (LONDON) LTD.
Berkeley Square House, London W1X 6BA

LIBRARY OF CONGRESS CATALOG CARD NUMBER: 70-162935

PRINTED IN THE UNITED STATES OF AMERICA

To my devoted wife, Ikuko

Contents

Preface xiii
Acknowledgments xv

Part I. **PRINCIPLES AND ELEMENTARY APPLICATIONS**

1. **Kinetic Theory**

1.1. BOLTZMANN EQUATION 3
1.2. MAXWELL–BOLTZMANN DISTRIBUTION FUNCTION 7
1.3. CALCULATION OF AVERAGES 9
1.4. SPECTRAL BROADENING BY THE DOPPLER EFFECT 11
1.5. MEAN FREE PATH 12
1.6. ELEMENTARY TREATMENT OF TRANSPORT PHENOMENA 13
1.7. BOLTZMANN AND GIBBS 17
REFERENCES 19

2. **Principles of Statistical Mechanics**

2.1. PHASE SPACE AND THE LIOUVILLE THEOREM 21
2.2.* ERGODIC THEORIES 23
2.3. *H*-THEOREM FOR SYSTEMS IN EQUILIBRIUM 30

* An asterisk beside a heading designates a specialized subject that can be omitted in a first reading.

2.4. MEANINGS OF THE CONSTANTS IN THE CANONICAL DISTRIBUTION FUNCTION 33
2.5. COARSE-GRAINING 36
2.6. PRODUCT APPROXIMATION FOR THE DISTRIBUTION FUNCTION 40
2.7. H-THEOREM BASED ON THE MASTER EQUATION 42
PROBLEMS 45
REFERENCES 47

3. Partition Functions

3.1. BOLTZMANN STATISTICS 48
3.2. PARTITION FUNCTION 51
3.3. GIBBS'S PARADOX 54
3.4. GRAND ENSEMBLE 57
3.5. RELATION BETWEEN THE CANONICAL AND GRAND CANONICAL PARTITION FUNCTIONS 59
3.6. FLUCTUATIONS 63
3.7. THE ELASTICITY OF RUBBER 65
3.8. LATTICE DEFECTS 68
PROBLEMS 69
REFERENCES 70

4. Ideal Bosons and Fermions

4.1. BLACKBODY RADIATION 71
4.2. SPECIFIC HEATS OF SOLIDS 75
4.3. QUANTUM STATISTICS OF IDEAL GASES 80
4.4. BOSE–EINSTEIN CONDENSATION 84
4.5. PHONONS AND ROTONS 89
4.6. HEAT CAPACITIES OF FERMI GASES AND FERMI LIQUIDS 91
4.7. ELEMENTARY TREATMENT OF TRANSPORT PHENOMENA IN DEGENERATE GASES 98
4.8. DE HAAS–VAN ALPHEN EFFECT 101
4.9.* PARASTATISTICS 105
PROBLEMS 107
REFERENCES 109

Part II. CLASSICAL INTERACTING SYSTEMS

5. Linked Cluster Expansion

5.1. SECOND VIRIAL COEFFICIENT 113
5.2. CLUSTER EXPANSION 120
5.3. VIRIAL EXPANSION 126
5.4. IRREDUCIBLE INTEGRALS 128

5.5. CUMULANT EXPANSION 131
5.6. RING DIAGRAM APPROXIMATION FOR A CLASSICAL ELECTRON GAS 134
5.7. THEORY OF CONDENSATION 136
5.8. POLARIZABLE GASES 140
5.9. BOUNDS OF THE FREE ENERGY 144
5.10.* CLUSTER EXPANSIONS FOR BINARY MIXTURES 148
PROBLEMS 150
REFERENCES 152

6. Distribution Functions

6.1. REDUCED LIOUVILLE EQUATION AND BOLTZMANN EQUATION 153
6.2. STRESS TENSOR IN NONEQUILIBRIUM FLUIDS 157
6.3. VISCOSITY COEFFICIENT OF FLUIDS 158
6.4. PLASMAS 161
6.5. VIRIAL EQUATION OF STATE 163
6.6. DETERMINATION OF FLUID STRUCTURE 165
6.7. CRITICAL OPALESCENCE 166
6.8. EXPANSIONS OF DISTRIBUTION FUNCTIONS 169
6.9.* NODAL EXPANSION 172
6.10.* HNC AND PY APPROXIMATIONS 174
6.11.* BORN–GREEN THEORY 176
PROBLEMS 179
REFERENCES 180

7. Brownian Motion

7.1. RANDOM WALKS AND BROWNIAN MOTION 183
7.2.* RANDOM WALKS ON LATTICES 186
7.3.* STOKES FRICTION AND EINSTEIN VISCOSITY 188
7.4. LANGEVIN'S EQUATION 192
7.5. FRICTION COEFFICIENT OF A BROWNIAN PARTICLE 195
7.6. AUTOCORRELATION FUNCTION 199
7.7. NEUTRON SCATTERING 202
7.8. THE FOKKER–PLANCK EQUATION 205
7.9.* SELF-AVOIDING WALK PROBLEM 207
PROBLEMS 211
REFERENCES 213

8. Lattice Statistics

8.1. ONE-DIMENSIONAL LATTICE 215
8.2.* HELIX-COIL TRANSITION IN POLYPEPTIDE AND "MELTING" OF DNA 217

8.3. DUALITY PRINCIPLE 219
8.4.* RIGOROUS THEORY OF A TWO-DIMENSIONAL RECTANGULAR LATTICE 222
8.5. SPIN CORRELATION FUNCTIONS 230
8.6. LATTICE GAS 232
8.7. DISTRIBUTION OF ZEROS OF THE GRAND PARTITION FUNCTION 236
8.8. FREQUENCY SPECTRUM 239
8.9.* LATTICE GREEN'S FUNCTION 244
8.10.* SPHERICAL MODEL 246
8.11.* HEISENBERG MODEL 247
REFERENCES 250

9. **Phenomena near the Critical Temperature**

9.1. CRITICAL TEMPERATURE OF A FLUID 252
9.2. RELATIONSHIPS AMONG THE CRITICAL EXPONENTS 256
9.3. MAGNETIC PHASE TRANSITIONS 257
9.4. BINARY MIXTURES 259
9.5. QUANTUM LIQUID SOLUTION 261
9.6. ORDER–DISORDER THEORY 263
9.7. DENSITY FLUCTUATIONS NEAR CRITICAL TEMPERATURE 266
9.8. SPATIAL CORRELATION OF A BOSE GAS NEAR THE CONDENSATION TEMPERATURE 270
9.9. TRANSPORT COEFFICIENTS NEAR CRITICAL POINTS 271
REFERENCES 272

Part III. QUANTUM INTERACTING SYSTEMS

10. **Propagator Methods for the Partition Functions**

10.1. DENSITY MATRIX 278
10.2. DENSITY MATRIX IN THE CANONICAL ENSEMBLE 280
10.3. SIMPLE EXAMPLES OF THE DENSITY MATRIX 283
10.4. PROPAGATOR IN THE r–β SPACE 286
10.5. GRAPHIC REPRESENTATION OF PROPAGATORS 290
10.6. LINKED CLUSTER EXPANSION OF THE EQUATION OF STATE 293
10.7. EQUATION OF STATE IN THE RING DIAGRAM APPROXIMATION 295
10.8.* THE EIGENVALUES OF QUANTUM PROPAGATORS 299
10.9. CORRELATION ENERGY OF AN ELECTRON GAS 304
PROBLEMS 307
REFERENCES 309

11. **Propagator Methods for Distribution Functions**

11.1. LINKED CLUSTER EXPANSIONS OF DISTRIBUTION FUNCTIONS 311
11.2. DISTRIBUTION FUNCTIONS OF IDEAL QUANTUM GASES 313
11.3. TRIPLET DISTRIBUTION FUNCTION 316
11.4. CHAIN DIAGRAM APPROXIMATION 317
11.5. CLASSICAL ELECTRON GAS 320
11.6.* CHARGED FERMIONS 324
11.7.* CHARGED BOSONS 327
11.8.* PHONON SPECTRUM AND SPATIAL CORRELATIONS IN A HARD-SPHERE BOSE GAS 328
11.9.* HARD-SPHERE FERMIONS 332
PROBLEMS 335
REFERENCES 337

12. **Transport Phenomena in Degenerate Systems**

12.1. UEHLING–UHLENBECK EQUATION 338
12.2. TRANSPORT COEFFICIENTS 341
12.3. TRANSPORT PHENOMENA IN DEGENERATE SYSTEMS 343
12.4. PHONON–PHONON SCATTERING 348
12.5. ELECTRICAL CONDUCTIVITY OF METALS 352
12.6.* RESISTANCE MINIMA IN DILUTE MAGNETIC ALLOYS 355
PROBLEMS 358
REFERENCES 359

13. **Irreversibility and Transport Coefficients**

13.1. RESPONSE TO EXTERNAL FORCES 361
13.2. KRAMERS–KRONIG RELATIONS FOR THE RESPONSE FUNCTION 364
13.3. SYMMETRY PROPERTIES OF RESPONSE FUNCTIONS 366
13.4. RESPONSE IN CANONICAL ENSEMBLES 367
13.5. TRANSPORT COEFFICIENTS 369
13.6. REDUCTION OF TRANSPORT COEFFICIENTS 372
13.7.* CLUSTER EXPANSION OF TIME CORRELATION FUNCTIONS 375
13.8.* MASTER EQUATION 376
PROBLEMS 379
REFERENCES 380

14. **Second Quantization**

14.1. NUMBER OPERATOR 383
14.2. INTERACTION HAMILTONIAN 385
14.3. LATTICE VIBRATIONS 386
14.4. PHONON SPECTRUM IN A DEGENERATE BOSE OAS 388

14.5. ELECTRON GAS 390
14.6. ELECTRON–PHONON INTERACTION 392
14.7. ELECTRON–ELECTRON INTERACTION VIA PHONONS 395
14.8. *f*-SUM RULE 397
14.9. DIELECTRIC CONSTANT OF A PLASMA 398
14.10.* SPIN AND STATISTICS 400
PROBLEMS 403
REFERENCES 404

15. **Green's Functions**

15.1. TEMPERATURE GREEN'S FUNCTION 405
15.2. PROPERTIES OF THE GREEN'S FUNCTION 407
15.3. CONTRACTION 409
15.4. PERTURBATION CALCULATION OF THE GRAND PARTITION FUNCTION 412
15.5. INTERACTION REPRESENTATION AND LEHMANN REPRESENTATION 416
15.6. APPLICATIONS OF A ONE-BODY TIME GREEN'S FUNCTION 418
15.7. RESPONSE FUNCTIONS 422
15.8. EQUATIONS OF MOTION 426
15.9.* SUPERCONDUCTIVITY 427
REFERENCES 432

Subject Index 433

Preface

The macroscopic properties of a physical system may be described by continuum physics, but if our interest is in explaining rather than describing the properties, their origin must be traced back to the constitutional particles. Unfortunately, for a system of many particles the equations of motion are in general not soluble and the initial conditions necessary for solving the equations are not known.

Statistical mechanics bypasses this difficulty by considering all possible states in which a system can be and by finding the probability for the realization of each state. A macroscopic quantity is obtained by evaluating a certain weighted average of a physical quantity over such states. In this way, atomic physics and macroscopic continuum physics are linked by statistical mechanics. It is one of the fundamental branches of theoretical physics and chemistry and deals with many systems such as gases, liquids, solids, and even molecules which have many atoms.

This book is written for a one-year graduate course in statistical mechanics or a half-year course followed by a half-year course on related subjects such as special topics and applications or elementary many-body theories. Efforts are made such that discussions of each subject start with an elementary level and end at an advanced level.

This book consists of three parts. Part I, which includes Chapters 1–4, gives the principles, with elementary applications to noninteracting systems. It begins with kinetic theory and discusses classical and quantum systems in

equilibrium and nonequilibrium. Examples in this part do not directly deal with interactions. Nevertheless, an overall picture of statistical mechanics may be obtained here at an elementary level. In Part II, which consists of Chapters 5–9, classical statistical mechanics is developed for interacting systems in equilibrium and nonequilibrium. Finally, in Part III, quantum statistics is presented to an extent which enables the reader to proceed to advanced many-body theories.

In teaching, the author has found that it is often insufficient to discuss a subject only once. For memory and enjoyment, recapitulations such as found in sonatas in music help students. Therefore, in the three parts, some subjects are discussed repeatedly, although on different levels. For example, important results on Brownian motion that are discussed in Chapter 7 reappear in different forms in later sections. The ring diagram approximation, first introduced in Section 5.6 for a classical electron gas, is used in Section 10.7 to derive an important formula for quantum gases. This formula is rederived in Section 15.4 by the Green's function method. The Maxwell transport theory introduced in Section 1.6 is given a quantum mechanical flavor in Section 4.7 and improved in later sections.

Most chapters contain a brief summary of the historical development of the subject, special topics, and problems. Some of the problems are not for simple class tests but rather for stimulation of further studies. Difficult problems, and sections with advanced or special topics are marked by asterisks.

Although selection of topics depends on the reader and the instructor, the author recommends the following:

1. *Thermal Physics and Introductory Stastical Mechanics* (*half-year*)
   Part I and selected sections from Part II such as 5.1, 6.1, 7.4, 7.5, 9.1, 9.2, etc.
2. *Statistical Mechanics* (*one-year*)
   Most of the subjects in Part II plus selected subjects from Parts I and III. Sections with asterisks may be omitted.
3. *Special Topics in Statistical Physics* (*half-year*)
   This course depends very much on instructors, but sections with asterisks may serve.
4. *Introductory Many-Body Theory* (*half-year*)
   Part III plus classical subjects such as those in 5.5, 5.6, 6.10.

Acknowledgments

Many parts of this book were based on the author's research supported by the National Science Foundation. The author is grateful to the investigators, the societies, and the publishers who have given him permission to use tables and diagrams. The author is indebted to Professors Lars Onsager and George E. Uhlenbeck for important comments. Special thanks are due to his colleagues, especially to Drs. Beth, Yeh, Wadati, and Godoy for reading the manuscript and correcting errors and to Professor Ernest M. Loebl and the staff of Academic Press for constant interest, encouragement, and cooperation.

PART

I

PRINCIPLES AND ELEMENTARY APPLICATIONS

1

Kinetic Theory

The foundation of kinetic theory was laid by Maxwell (1) and Boltzmann (2) approximately one hundred years ago. Maxwell published his famous memoire in 1866 and Boltzmann introduced his H-theorem in 1872. In recognition of their contributions, we begin this book with Boltzmann's equation and his H-theorem, and applications of the Maxwell–Boltzmann velocity distribution function to equilibrium and nonequilibrium gases.

Maxwell evaluated the viscosity and the thermal conductivity of a dilute gas. His results were surprising to many of his contemporaries, because the temperature and density dependences of these coefficients were not what they expected. Following Maxwell, we shall try to find answers to questions such as: "How does the viscosity of a dilute gas change with temperature and density?"

From the rigorous statistical mechanical point of view, the Maxwell theory is not correct, although it does give a roughly correct representation of viscosity and thermal conductivity.

Statistical mechanics can be studied without going through kinetic theory. However, because of its simplicity, kinetic theory is instructive and helpful for the study of statistical mechanics.

1.1. Boltzmann Equation

We shall consider a gas that is composed of N identical molecules in a vessel of volume V at temperature T. We shall assume that the system is in equilibrium and that the molecules are spherical and perfectly elastic. In their

collisions, the total momentum and energy are conserved. The density of the gas is considered low, so that the effects of other molecules are negligible when two molecules are colliding. The gas is almost ideal.

Let us denote by $f(v_x, v_y, v_z)\, dv_x\, dv_y\, dv_z$ the number of gas molecules with velocities between v_x and $v_x + dv_x$, and v_y and $v_y + dv_y$, v_z and $v_z + dv_z$ in the x, y, and z directions, respectively. We shall denote this number by $f(\mathbf{v})\, d\mathbf{v}$. The function $f(v_x, v_y, v_z)$, or $f(\mathbf{v})$, is called the *velocity distribution function*.

In this section, we shall consider homogeneous and isotropic gases with no external forces acting. Then, the distribution function is independent of coordinates, and depends only on v, the magnitude of velocity $\mathbf{v}$. Therefore, the distribution function may be expressed as $f(v^2)$. The function $f(\mathbf{v})$ must satisfy the normalization condition

$$\int_{-\infty}^{\infty} f(\mathbf{v})\, d\mathbf{v} = N. \tag{1.1}$$

Here, the integration is extended over all values of velocity components v_x, v_y, and v_z, namely from $-\infty$ to ∞. Of course, there are no molecules with infinite velocity. This fact requires that $f(\mathbf{v})$ must vanish in the limit $v \to \infty$.

Collisions among the gas particles may result in a change in the velocity distribution function. Following Boltzmann, the change may be evaluated under the *collision hypothesis* (*Stosszahlansatz*). In this hypothesis, it is assumed that colliding particles are distributed at random without correlation between velocity and position. Thus, the number of collisions between a pair of particles of specified velocities is evaluated by applying a velocity distribution function in a microscopic *collision cylinder*. Whether or not this procedure is legitimate is a matter subject to question. At least it may not be appropriate if the gas is not in a state of *molecular chaos*.

Let us find a collisional change in $f(\mathbf{v})$ by considering a representative collision of a pair of particles of velocities $\mathbf{v}$ and $\mathbf{v}_1$. After collision, the velocities become $\mathbf{v}'$ and $\mathbf{v}_1'$, respectively. The particles are assumed to be perfectly elastic spheres of diameter a. As is illustrated in Fig 1.1, one of the particles approaches the target particle with a relative velocity $\mathbf{g} = \mathbf{v}_1 - \mathbf{v}$ and leaves with $\mathbf{g}' = \mathbf{v}_1' - \mathbf{v}'$. We specify the direction of collision by a unit vector $\mathbf{n}$ of the external bisector of the angle between $\mathbf{g}$ and $\mathbf{g}'$. We find

$$\mathbf{v}' = \mathbf{v} - \mathbf{n}(\mathbf{n}\cdot\mathbf{g}), \qquad \mathbf{v}_1' = \mathbf{v}_1 + \mathbf{n}(\mathbf{n}\cdot\mathbf{g}), \qquad |\mathbf{g}| = |\mathbf{g}'|. \tag{1.2}$$

Evaluating the Jacobian for given $\mathbf{n}$ and $\mathbf{g}$, we find

$$d\mathbf{v}\, d\mathbf{v}_1 = d\mathbf{v}'\, d\mathbf{v}_1'. \tag{1.3}$$

Note that in this expression the two velocities are interrelated. According to Eq. (1.3), the elementary volume in the velocity space is invariant with respect to collision.

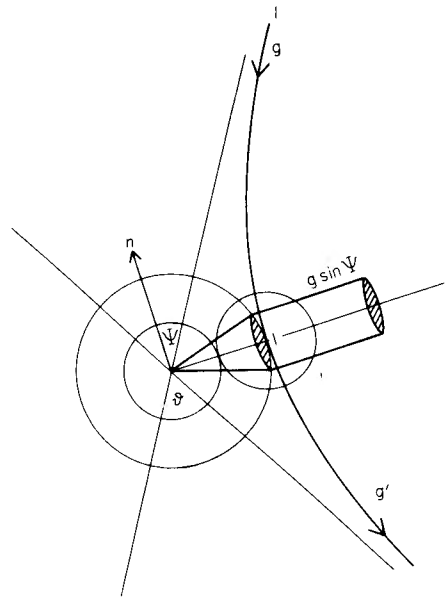


Fig. 1.1. Geometry of a two-particle encounter.

On the surface of a sphere of radius a that is concentric to the target sphere, we now construct a collision cylinder. The particle 1 has a velocity component $-g \sin \Psi$ along the direction connecting the centers of the two molecules (Fig. 1.1). Let us introduce a cylinder of side length $1 \sec \times g \sin \Psi$ and of base $a^2 \sin(\frac{1}{2}\pi - \Psi)\, d\Psi\, d\Phi$. All the particles with velocities between $\mathbf{v}_1$ and $\mathbf{v}_1 + d\mathbf{v}_1$ in this cylinder will collide with the target particle in unit time. Their number is $f(\mathbf{v}_1)\, d\mathbf{v}_1$ times the volume of the collision cylinder:

$$f(\mathbf{v}_1)\, d\mathbf{v}_1 (g \sin \Psi) a^2 \sin(\tfrac{1}{2}\pi - \Psi)\, d\Psi\, d\Phi = f(\mathbf{v}_1)\, d\mathbf{v}_1\, g I(\vartheta, g)\, d\Omega,$$

where $d\Omega = \sin\vartheta\, d\vartheta\, d\Phi$ and $\pi I(\vartheta, g)$ is the cross section, which, in our case, is equal to $\pi a^2/4$. Here, ϑ is the scattering angle.

There are $f(\mathbf{v})\, d\mathbf{v}$ similar target molecules with velocities between $\mathbf{v}$ and $\mathbf{v} + d\mathbf{v}$. Therefore, the number of collisions of the above type per unit time is given by

$$A = f(\mathbf{v}) f(\mathbf{v}_1)\, d\mathbf{v}\, d\mathbf{v}_1 g I(\vartheta, g)\, d\Omega. \tag{1.4a}$$

There exists an inverse to the above type of collision: The two colliding particles can have velocities $\mathbf{v}'$, $\mathbf{v}_1'$ before collision and $\mathbf{v}$, $\mathbf{v}_1$ after collision. Particle 1 approaches the target with relative velocity $\mathbf{g}'$ and leaves with velocity $\mathbf{g}$. Thus, the direction of collision specified by $\mathbf{n}$ is just the opposite, but the collision is still determined by the same scattering angle ϑ. There are

$f(\mathbf{v}')\,d\mathbf{v}'$ target molecules, and the total number of inverse collisions per unit time is

$$B = f(\mathbf{v}')f(\mathbf{v}_1')\,d\mathbf{v}'\,d\mathbf{v}_1'\,gI(\vartheta, g)\,d\Omega. \tag{1.4b}$$

The net change, which is given by the difference $B - A$ integrated over $\mathbf{v}_1$ or $\mathbf{v}_1'$ and the angles, may be denoted by $(df/dt)\,d\mathbf{v}$. Using Eq. (1.3), we find

$$\frac{df}{dt} = \int (f'f_1' - ff_1)gI(\vartheta, g)\,d\mathbf{v}_1\,d\Omega. \tag{1.5}$$

This is called the *Boltzmann equation*. In the integrand, f, f_1, etc. mean $f(\mathbf{v})$, $f(\mathbf{v}_1)$, etc., respectively.

Using Eq. (1.5), Boltzmann proved the famous *H-Theorem* in 1872. Let us introduce Boltzmann's H-function, defined by

$$H = \int f \log f\,d\mathbf{v}. \tag{1.6}$$

By making use of the Boltzmann equation, the change in H may be obtained as follows:

$$\frac{dH}{dt} = \int \frac{df}{dt}(\log f + 1)\,d\mathbf{v}$$

$$= \int (f'f_1' - ff_1)(\log f + 1)gI(\vartheta, g)\,d\mathbf{v}\,d\mathbf{v}_1\,d\Omega. \tag{1.7}$$

Here, $\mathbf{v}$ and $\mathbf{v}_1$ are dummy variables. Since the integrations are from $-\infty$ to ∞, we can switch $\mathbf{v}$ and $\mathbf{v}_1$. The result is

$$\frac{dH}{dt} = \int (f_1'f' - f_1 f)(\log f_1 + 1)gI(\vartheta, g)\,d\mathbf{v}_1\,d\mathbf{v}\,d\Omega.$$

One can also switch $\mathbf{v}$ and $\mathbf{v}'$ and/or $\mathbf{v}_1$ and $\mathbf{v}_1'$. In these replacements, the differential volume remains the same because of Eq. (1.3). Adding the four expressions thus obtained, we obtain

$$\frac{dH}{dt} = -\frac{1}{4}\int (f'f_1' - ff_1)\log\left(\frac{f'f_1'}{ff_1}\right)gI(\vartheta, g)\,d\mathbf{v}\,d\mathbf{v}_1\,d\Omega. \tag{1.7'}$$

One can easily see that the integrand is always nonnegative. Thus, we arrive at the Boltzmann H-theorem:

$$\frac{dH}{dt} \leqslant 0. \tag{1.8}$$

That is, the H-function decreases with time or stays constant. When it stops changing with time, the system is considered to have reached an equilibrium state. The system will be in equilibrium when the *detailed balance condition*

$$f'f_1' = ff_1 \tag{1.9}$$

is satisfied. This condition means that the product $f(\mathbf{v})f(\mathbf{v}_1)$ is invariant with respect to collision and should be a function of the constants of motion. This conclusion leads to the Maxwell–Boltzmann distribution function to be discussed shortly, and the Maxwell–Boltzmann distribution function satisfies $dH/dt = 0$. Boltzmann concluded that the Maxwell–Boltzmann distribution is necessary and sufficient for equilibrium.

1.2. Maxwell–Boltzmann Distribution Function

Equation (1.9) has been derived under the assumption that the collisions of two molecules are not disturbed by any third molecule or by the container.

Elastic collisions of two molecules can be described in a symmetric way in a coordinate system fixed to the center of gravity. Therefore, any proportionality factors which might have been necessary for ff_1 and $f'f_1'$ in Eqs. (1.4a) and (1.4b) must be the same. Thus, we have arrived at the detailed balance condition (1.9).

The distribution function $f(\mathbf{v})$ depends only on the absolute magnitudes of velocities. Let us use Eq. (1.9) for the particular set of velocities

$$\mathbf{v}' = 0, \qquad v_1'^2 = v^2 + v_1^2.$$

Then, the detailed balance condition gives

$$f(0)f(v^2 + v_1^2) = f(v^2)f(v_1^2).$$

Differentiating both sides with respect to v_1^2 and setting $v_1 = 0$, we obtain

$$\frac{\partial f(v^2)}{\partial v^2} = -\alpha f(v^2), \tag{2.1}$$

where

$$\alpha = -\frac{1}{f(0)}\left[\frac{\partial f(v_1^2)}{\partial v_1^2}\right]_{v_1 = 0}.$$

Equation (2.1) is the differential equation to determine $f(v^2)$, from which we find, upon integration,

$$f(v^2) = A\exp(-\alpha v^2). \tag{2.2}$$

The integration constant A and the unknown constant α can be determined

by the energy conservation law and by the requirement that the total number of molecules is N. The latter condition is expressed in Eq. (1.1) and gives

$$A = N\left(\frac{\alpha}{\pi}\right)^{3/2}. \tag{2.3}$$

The mean kinetic energy U for the total system is

$$\begin{aligned} U &= \int_{-\infty}^{\infty} \tfrac{1}{2}mv^2 f(\mathbf{v})\, d\mathbf{v} \\ &= 2\pi m N\left(\frac{\alpha}{\pi}\right)^{3/2} \int_0^{\infty} v^4 e^{-\alpha v^2}\, dv = \frac{3mN}{4\alpha}, \end{aligned} \tag{2.4}$$

where m is the mass of a gas molecule. This equation shows that the unknown constant α is inversely proportional to the mean kinetic energy per particle. A more convenient expression for α can be derived if we calculate the pressure.

The pressure p of a gas is given by the momentum change per unit time per unit area of the gas molecules colliding against the wall of vessel. The momentum change per unit time is a force, and force per unit area of surface is pressure. Let us consider collisions of gas molecules with velocities $\mathbf{v}$ and $\mathbf{v} + d\mathbf{v}$ with the wall that is perpendicular to the x axis. The momentum change in one collision is $2mv_x$ and the number of molecules colliding with the wall in 1 sec is $(v_x/V)f(\mathbf{v})\, d\mathbf{v}$. If the volume of vessel is V and if the system is isotropic and free from interaction, the pressure is

$$\begin{aligned} p &= 2\frac{m}{V} \iiint_{-\infty, v_x>0}^{\infty} v_x^2 f(\mathbf{v})\, d\mathbf{v} \\ &= \frac{2}{3}\frac{1}{2}\frac{m}{V} \iiint_{-\infty}^{\infty} v^2 f(\mathbf{v})\, dv_x\, dv_y\, dv_z. \end{aligned} \tag{2.5}$$

On the other hand, for an ideal gas, we have the equation of state

$$pV = NkT, \tag{2.6}$$

where k is the Boltzmann constant and T is the absolute temperature. Comparing Eqs. (2.5), (2.6), and (2.4), we have

$$U = \tfrac{3}{2}NkT = \tfrac{3}{2}pV. \tag{2.7}$$

Therefore, the unknown constant α is given by

$$\alpha = \frac{m}{2kT}. \tag{2.8}$$

From this equation, we see that α is a function of the mass of the molecules and the temperature and has the dimension of reciprocal velocity squared.

Introducing Eqs. (2.3) and (2.8) into Eq. (2.2), we obtain

$$f(\mathbf{v}) = N\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left(-\frac{mv^2}{2kT}\right), \tag{2.9}$$

which is known as the *Maxwell–Boltzmann velocity distribution function*. The exponential function is determined by the dimensionless ratio of the kinetic energy to the thermal energy kT. Boltzmann showed that, if the gas molecules have a potential energy in addition to the kinetic energy, the exponential factor is determined by the total energy, namely the kinetic energy plus the potential energy, divided by kT.

The above distribution function was first derived by Maxwell in 1859. However, his derivation was based on the assumption that the velocity distribution of gas molecules in one direction is independent of those in the other directions. Moreover, the molecular collisions were not taken into consideration in his theory.

Although Maxwell's assumptions are plausible for a homogeneous gas, these had yet to be justified. Boltzmann improved Maxwell's theory in this respect. The above proof is based on Boltzmann's method. Equation (2.9) represents the Maxwell–Boltzmann velocity distribution law.

In the case when the gas moves as a whole with a velocity $\mathbf{u}$, the distribution function assumes the following form:

$$f(\mathbf{v}; \mathbf{u}) = N\left(\frac{m}{2\pi kT}\right)^{3/2} \exp\left[-\frac{m(\mathbf{v}-\mathbf{u})^2}{2kT}\right]. \tag{2.10}$$

This expression can be proved by using energy and momentum conservation.

1.3. Calculation of Averages

The Maxwell–Boltzmann distribution function enables us to calculate the averages of functions of velocities. The simplest example is the velocity itself, but its average is zero if there is no macroscopic flow. The distribution function is an even function of velocity and averages of odd functions of velocity should be zero. However, there exists an average speed. If we denote this by $\bar{v}$, then

$$\bar{v} = \frac{1}{N}\int_0^{\infty} v\, \tilde{f}(v)\, dv = 2\left(\frac{2kT}{m\pi}\right)^{1/2} = 1.13\left(\frac{2kT}{m}\right)^{1/2}. \tag{3.1}$$

Instead of the distribution function $f(\mathbf{v})$, which gives the number of molecules of velocities between $\mathbf{v}$ and $\mathbf{v} + d\mathbf{v}$, the distribution function $\tilde{f}(v)$ gives the number of molecules having speeds between v and $v + dv$:

$$\tilde{f}(\mathbf{v}) = 4\pi v^2 f(\mathbf{v}). \tag{3.2}$$

$f(\mathbf{v})$ decreases monotonically as the velocity increases. On the other hand, $\tilde{f}(v)$ has a maximum at the speed given by

$$v_m = \left(\frac{2kT}{m}\right)^{1/2}. \tag{3.3}$$

v_m is the most probable speed. It is interesting to observe that the average speed $\bar{v}$ is 1.13 times larger than v_m. By using the Maxwell–Boltzmann distribution function, one can show that the molecules of speed greater than $\bar{v}$ comprise only 46.7% of all molecules, that is, less than half of the total number of molecules.

Similarly, we can calculate the average square velocity as follows:

$$\overline{v^2} = \frac{1}{N}\int_{-\infty}^{\infty} v^2 f(\mathbf{v})\, d\mathbf{v} = \frac{3}{2}\left(\frac{2kT}{m}\right). \tag{3.4}$$

Since the system is assumed homogeneous, $\overline{v_x^2} = \overline{v_y^2} = \overline{v_z^2} = \overline{v^2}/3$ and the average kinetic energies in the x, y, and z directions are the same and equal to $kT/2$. This result is an example of the equipartition of energy.

Near room temperature, the average speed of a gas molecule as given by Eq. (3.1) is not small. For instance, hydrogen gas molecules travel at 1750 m sec^{-1} at 273°K, and oxygen molecules travel at 425 m sec^{-1} at the same temperature. On the other hand, the diffusion velocities of these gas molecules are less than the above values because the collisions of molecules moving in all directions determine the diffusion.

Another important average quantity is the number z_1 of molecules that collide with a unit area of wall per second. Taking the x axis perpendicular to the wall, z_1 is given by

$$z_1 = \frac{1}{V}\int_0^{\infty}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} v_x f(\mathbf{v})\, dv_x\, dv_y\, dv_z$$

$$= \frac{N}{V}\left(\frac{kT}{2\pi m}\right)^{1/2} = \frac{1}{4}\frac{N}{V}\bar{v}, \tag{3.5}$$

where V is the total volume. For a pressure of 760 Torr and a temperature of 300°K, z_1 is 1.1×10^{24} in the case of hydrogen gas. That is, 1.8 moles of the hydrogen gas collide with a unit area of the wall per unit time.

Problem

If ρ is the density and $(\overline{v_x^2})^{1/2}$ is equal to $(p/\rho)^{1/2}$, which is Newton's expression for the speed of sound. Why is this not the actual speed of sound?

1.4. Spectral Broadening by the Doppler Effect

As shown in the previous section, various properties of gases are related to certain averages of velocity. The averages are determined by the distribution function, but the distribution function itself cannot be determined by knowing only such averages. The spectral broadening due to the Doppler effect is determined directly by the distribution function and allows us to examine experimentally the form of the function.

As is well known, gases are characterized by so-called line spectra. In addition to a natural linewidth, there are several effects which broaden spectral lines: For instance, collisions among molecules and collisions between molecules and the wall cause broadening. So does the Doppler effect.

Of these effects, collisions among molecules give a large contribution that is proportional to the pressure and is called *pressure broadening*. Collisions between molecules and the wall are important only when the pressure is very low and the mean free path of the molecules, which will be discussed in the next section, becomes comparable to the linear dimensions of the vessel. In the normal situation, pressure broadening is more important than the Doppler effect. However, the latter is not proportional to the pressure and can be distinguished from pressure broadening.

If a gas molecule emitting light of wavelength λ moves away from an observer, the wavelength measured by the observer is longer due to the Doppler effect and is given by (for $c \gg u$)

$$\lambda = \lambda_0\left(1 + \frac{u}{c}\right),$$

where c is the velocity of light. A similar formula holds for the case when the molecule approaches the observer. Generally, u is the velocity component in the direction of observation.

If there are many molecules that emit light of wavelength λ while moving with velocity component between u and $u + du$, the observed wavelength can have the width

$$d\lambda = \lambda_0 \frac{du}{c}.$$

The intensity of the observed light $I(\lambda)\,d\lambda$ must be proportional to the number $f(u)\,du$ of molecules of velocity between u and $u + du$. Since $u = c(\lambda - \lambda_0)/\lambda_0$, we have

$$I = I_0 \exp\left[\frac{-mc^2(\lambda - \lambda_0)^2}{2\lambda_0^2\, kT}\right], \tag{4.1}$$

where I_0 is a proportionality constant. Equation (4.1) gives the intensity distribution, which is simply the Maxwell–Boltzmann distribution. According to Eq. (4.1), the logarithm of intensity plotted against $(\lambda - \lambda_0)^2$ is a straight line of slope $-mc^2/2\lambda_0^2 kT$. This conclusion was confirmed by Fabry and Ruisson in 1912 and more accurately by Ornstein and Van Wyk in 1932.

1.5. Mean Free Path

The Maxwell–Boltzmann distribution function was derived by considering molecular collisions. The distance traveled by a molecule between two successive collisions is called the *free path* and its average is called the *mean free path.*

The mean free path l is equal to the average speed $\bar{v}$ divided by the number of collisions in 1 sec. Suppose that a molecule of diameter a collides with another molecule at rest. Those molecules that are contained in a cylinder of length $\bar{v} \times 1$ sec and with a circular base of radius a will collide with the molecule at rest. Therefore, in 1 sec, this molecule experiences $\pi a^2 n\bar{v}$ collisions, where n is the number density. Thus, the mean free path is given by

$$l = \frac{1}{\pi a^2 n}. \tag{5.1}$$

Under standard conditions (0°C, 1 atm), n is equal to Loschmidt's number, 2.687×10^{19}, and l is of the order of 10^{-6} cm for $a = 5$ Å.

A more rigorous expression for l may be obtained by taking for the side length of the cylinder the magnitude of the relative velocity (multiplied by 1 sec) in the direction of the line connecting the centers of the colliding molecules, because actually both molecules are in motion. Let us consider molecules with velocities $\mathbf{v}_1$ and $\mathbf{v}_2$. The relative velocity $\mathbf{v}_2 - \mathbf{v}_1$ is denoted by $\mathbf{g}$. Per unit volume, the numbers of molecules with these velocities are given by $(1/V)f(\mathbf{v}_1)\,d\mathbf{v}_1$ and $(1/V)f(\mathbf{v}_2)\,d\mathbf{v}_2$. As can be seen from Fig. 1.1, we count the molecules in the cylinder of volume $(a^2 g/4)\,d\Omega$, with $d\Omega = \sin\vartheta\, d\vartheta\, d\Phi$. The number z_2 of collisions per second per volume is

$$z_2 = \frac{1}{2V^2}\iint a^2 f(\mathbf{v}_1) f(\mathbf{v}_2)\frac{g}{4}\,d\Omega\, d\mathbf{v}_1\, d\mathbf{v}_2. \tag{5.2}$$

The integration is carried out over all velocity range of $\mathbf{v}_1$ and $\mathbf{v}_2$, the magnitude of the relative velocity being g. Equation (5.2) may be calculated by using the Maxwell–Boltzmann distribution function. Remembering that a collision gives an end point of two free paths, a formula for the mean free path which improves the one given by Eq. (5.1) can be obtained by taking the ratio of $n\bar{v}$ to $2z_2$. Thus,

$$l = \frac{1}{\pi n a^2\sqrt{2}}. \tag{5.3}$$

In comparison to Eq. (5.1), Eq. (5.3) differs only by a factor $1/\sqrt{2}$.

1.6. Elementary Treatment of Transport Phenomena

Let us consider a gas with a nonuniform distribution of temperature or average velocity. That is, one part of the gas may have higher energy or higher momentum compared to another part. On the average, molecules from such a part will lose some of their energy or momentum in colliding with molecules from another part. In other words, there is a transport of energy or of momentum; such transport causes viscosity or thermal conductivity, respectively. Irreversible processes due to such transport are called *transport phenomena.*

Let us assume that a certain property changes along the z direction. Consider a plane surface A of unit area which is at some value of z perpendicular to the z axis (Fig. 1.2). The number of molecules with velocities between $\mathbf{v}$

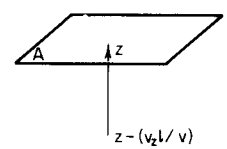


Fig. 1.2. Transport across a plane.

and $\mathbf{v} + d\mathbf{v}$ that pass the surface A in 1 sec is $v_z f(\mathbf{v})\,d\mathbf{v}/V$. The molecules at z carry some physical quantity $G(z)$. If the mean free path is l, the molecules at $z - (v_z l/v)$ will pass the surface A in 1 sec. These molecules carry the quantity $G[z - (v_z l/v)]$. Generally, since l is small compared with the change in macroscopic scale z, we make the approximation

$$G\left(z - \frac{v_z l}{v}\right) = G(z) - \frac{v_z l}{v}\frac{\partial G}{\partial z}.$$

The total quantity $\Gamma(z)$ transported is

$$\Gamma(z) = G(z) \int \frac{f(\mathbf{v})}{V} v_z \, d\mathbf{v} - l \frac{\partial G}{\partial z} \int v_z^2 \frac{f(\mathbf{v})}{vV} \, d\mathbf{v}.$$

The integration over $v_z > 0$ ($v_z < 0$) corresponds to the contribution of molecules passing through the surface A upward (downward). The transported quantity $\Gamma(z)$ is the difference between these contributions, and the integration is carried over all the velocity space. The first integral in the above equation vanishes, and integration of the second gives

$$\Gamma(z) = -\frac{1}{3} l \frac{\partial G}{\partial z} n\bar{v}. \tag{6.1}$$

The minus sign indicates that the transport takes place in the direction from higher to lower G.

1.6.1. Momentum Transport

Let us consider the case where the gas flows parallel to the xy plane with the flow velocity u dependent on z. The quantity transported is the momenta of gas molecules, and we put $G = mu$. The momentum change per unit time equals the force F; for this case, Γ in Eq. (6.1) is F. On the other hand, the force is proportional to the velocity gradient, with viscosity as the proportionality constant:

$$F = -\eta \frac{\partial u}{\partial z}. \qquad \text{(per area)} \tag{6.2}$$

Combining Eqs. (6.2) and (6.1), we obtain

$$\eta = \tfrac{1}{3} l n m \bar{v}. \tag{6.3}$$

If we introduce for $\bar{v}$ and l of Eq. (6.3) the expressions given in Eqs. (3.1) and (5.3), we have

$$\eta = \frac{2}{3\pi a^2} \left(\frac{mkT}{\pi}\right)^{1/2}. \tag{6.4}$$

This shows that the viscosity η depends on the temperature and is independent of the density or the pressure. This result, when first derived by Maxwell, was considered strange. However, it was confirmed by later experiments, giving support to the kinetic theory of gases. As we can see from Eq. (6.3), the density independence comes from the fact that the mean free path is inversely proportional to the gas density.

For very low densities, where the mean free path becomes of the order of the linear size of the vessel, and for densities so high that liquefaction takes place, we expect deviations from Eq. (6.4). For example, the viscosity of CO_2 gas depends roughly on the density squared for temperatures 20–40°C and for pressure 60–100 atm.

According to Eq. (6.4), the viscosity of gases is proportional to $\sqrt{T}$. Experimentally, the temperature dependence is slightly stronger. This is ascribed to molecular forces.

1.6.2. Energy Transport

We consider the case where there is a temperature gradient in the z direction. The average energy per molecule U/N is a function of z. We choose $G(z) = U(z)/N$. Then,

$$\frac{dG}{dz} = \frac{1}{N} \frac{dU}{dT} \frac{dT}{dz} = mc_V \frac{dT}{dz},$$

where m is the molecular mass and c_V is the specific heat per atomic mass at constant volume. The energy transport becomes

$$\Gamma(z) = -\frac{1}{3} l n m c_V \bar{v} \frac{dT}{dz}.$$

On the other hand, the energy flow is connected to the thermal conductivity κ by the relation

$$\Gamma(z) = -\kappa \frac{dT}{dz}. \tag{6.5}$$

Therefore,

$$\kappa = \tfrac{1}{3} l n m c_V \bar{v}. \tag{6.6}$$

Comparing this with Eq. (6.3), we find

$$\frac{\kappa}{\eta} = c_V. \qquad \text{(per atomic mass)} \tag{6.7}$$

Table 1.1 shows that the observed values for $\kappa/\eta c_V$ are not equal to 1, contrary to the prediction of Eq. (6.7). Generally, $\kappa/\eta c_V$ is close to 2.5 for

TABLE 1.1

Viscosity and Thermal Conductivity of Gases

	$\eta \times 10^7$ (cgs)	$\kappa \times 10^3$ (cal cm^{-1} sec^{-1} deg^{-1})	γ	$\frac{\kappa}{\eta c_V}$	$\frac{(9\gamma - 5)}{4}$
He	1875	0.344	1.660	2.44	2.485
H_2	840	0.416	1.41	2.06	1.92
CO_2	1377	0.034	1.31	1.64	1.70
CH_4	1027	0.072	1.31	1.75	1.70

monatomic gases and is between 1 and 2.5 for polyatomic gases. The reason for this discrepancy is that the Maxwell–Boltzmann distribution is not the actual distribution if the gas is not homogeneous. By a more rigorous theory which takes into consideration the deviation of the distribution function from the Maxwell distribution, we obtain $\kappa/\eta c_V = 5/2$. For polyatomic gases, mc_V of Eq. (6.7) may be approximated by $(5/2)(3k/2) + c_i$, where c_i is the specific heat corresponding to the internal degrees of freedom: $c_i = mc_V - (3k/2)$. Use of $k = (\gamma - 1)mc_V$ gives $\kappa/\eta c_V = (9\gamma - 5)/4$, where γ is the ratio of the specific heat at constant pressure to that at constant volume. Table 1.1 shows that this relation is approximately true.

1.6.3. Diffusion of Molecules

If there exists a concentration gradient in the z direction, the molecules will move, on the average, from a region of higher concentration to one of lower concentration. Our problem is to find the transport of the molecules themselves. The molecules transport the number density n and instead of Eq. (6.1) we have

$$\Gamma(z) = -\frac{1}{3}\bar{v}l\frac{\partial n}{\partial z}, \tag{6.8}$$

where $\Gamma(z)$ is the quantity transported by diffusion and is equal to $-D\,\partial n/\partial z$, where D denotes the diffusion constant. Equation (6.8) gives

$$D = \tfrac{1}{3}l\bar{v}. \tag{6.9}$$

This is the formula for self-diffusion. Usually, the order of magnitude of l is 10^{-6} cm and that of $\bar{v}$ is 10^3 cm sec^{-1}. Therefore, the order of magnitude of D is 10^{-3} cm^2 sec^{-1}.

If there are two different types of gas molecules (denoted by 1 and 2), the constant of mutual diffusion is

$$D_{12} = \frac{n_1 l_2 \bar{v}_2 + n_2 l_1 \bar{v}_1}{3(n_1 + n_2)}, \tag{6.10}$$

where n_1, l_1, and $\bar{v}_1$ are, respectively, the density, the mean free path, and the mean velocity for Type 1 molecules, and the corresponding definitions hold for Type 2 molecules.

Problem

Prove that the viscosity for a two-dimensional gas is

$$\eta = \tfrac{1}{2}\rho\bar{v}l, \tag{6.11}$$

where ρ is the density.

1.7. Boltzmann and Gibbs

At this point, it may be of interest to see the transition from kinetic theory to statistical mechanics by looking briefly at the lives of two key figures, Boltzmann and Gibbs, who laid the foundation of statistical mechanics.

Ludwig Boltzmann was born in 1844, when the kinetic concept of matter was still not well established. He played a most important role in establishing kinetic theory and starting statistical mechanics.

Influenced by Maxwell's kinetic theory, especially by his transport theory, Boltzmann tried to prove that the Maxwell velocity distribution function is the only one for gases in equilibrium. For this purpose, he started from a non-equilibrium gas and derived his famous integrodifferential equation. Not only was he able to show by this equation that the Maxwell distribution, once established, is not altered by collisions, but he also arrived at a very important result: A quantity called H must always decrease or remain constant. In other words, Boltzmann concluded that his H decreases steadily toward equilibrium, at which the Maxwell distribution holds.

Boltzmann's H-theorem immediately met strong opposition. For instance, Loschmidt argued that the H-theorem contradicted Newtonian mechanics. If one sequence of molecular motion decreases H, then, by reversing all the velocities of the molecules (dynamical reversibility), one should find a sequence that increases H. Zermelo also objected to the H-theorem because it seemed to contradict the recurrence theorem of Poincaré. To these objections, Boltzmann answered that his H decreases statistically rather than dynamically. In other words, if a gas is not in thermal equilibrium, it will *most likely* tend to reach equilibrium, and once an equilibrium state is reached, it will *almost always* stay there.

Boltzmann wrote a noteworthy book, "Vorlesungen über Gastheorie"(*3*). This book, consisting of Part I, dated 1896, and Part II, dated 1898, starts

with many challenges and much ambition, as can be seen from his foreword:

> . . . even the complicated formulas of Maxwell's theory of electromagnetism were often considered useless before Hertz's experiments. I hope this will not also be the general opinion concerning gas theory! [p. 22].

However, it ends with growing pessimism:

> . . . it will be a great tragedy for science if the theory of gases were temporarily thrown into oblivion because of momentary hostile attitudes towards it. . . . I am conscious of being only one individual struggling weakly against the current of time. But it still remains in my power to make a contribution in such a way that, when the theory of gases is again revived, not too much will have to be rediscovered [p 216].

Boltzmann's growing pessimism led him to commit suicide in 1906 —one of the greatest tragedies in the history of physics. Ironically, soon after his death, kinetic theory and his statistical concepts were accepted.

A complete history of Boltzmann's contributions to statistical mechanics would require a much more extensive discussion than is possible here. Hence, we mention only some of them: He generalized the Boltzmann equation to discrete cases, and treated the dissociation rate of molecules; he succeeded in deriving hydrodynamic equations for simple cases; he introduced statistical considerations in the Liouville theorem; extending the early kinetic theory, he explained the specific heat of polyatomic gases; he developed the Clausius virial theorem for statistical applications; and he even treated the gas–liquid condensation based on the van der Waals gas model.

Josiah Willard Gibbs was much more fortunate than Boltzmann (*4*). He was born in New Haven, Connecticut, in February 1839, and entered Yale College in 1854. He graduated from Yale in 1858, and after spending the next five years in research, he received a degree and became a tutor of Latin and natural philosophy there. At the end of this three-year term of appointment, he traveled to France and Germany, where he was influenced by Magnus at Berlin, and Kirchhoff and Helmholtz at Heidelberg. He stayed in Europe for three years; returning to the United States, he obtained a professorship at Yale, a position he held until his death in 1903.

Gibbs started making scientific contributions at the age of 34. His scientific approach is well characterized by his first two papers: "Graphical Methods in Thermodynamics of Fluids" and "A Method of Geometrical Representation of the Thermodynamic Properties of Substances by Means of Surfaces." These papers were followed in 1856 and 1878 by two parts of the famous paper, "On the Equilibrium of Heterogeneous Substances," in which he introduced the famous phase rule.

Gibbs contributed to vector analysis and to the electromagnetic theory of light. His last work, "Elementary Principles in Statistical Mechanics," is most important. With this book, he achieved for statistical mechanics what Maxwell had for electrodynamics with his "Treatise." The book was published in 1902, whereby a new branch of physics called statistical mechanics was born. What Gibbs did was to construct statistical mechanics based on the new concept of ensembles instead of the velocity space due to Maxwell and Boltzmann.

REFERENCES

1. J. C. Maxwell, *Philos. Trans. Roy. Soc. London, Ser. A* **CLVII**, 26 (1866).
2. L. Boltzmann, *Wein. Ber.* **66**, 275 (1872).
3. L. Boltzmann, "Lectures on Gas Theory" (translated by S. G. Brush). Univ. of California Press, Berkeley, 1964.
4. J. W. Gibbs, "Collected Works." Yale Univ. Press, New Haven, 1948.

2

Principles of Statistical Mechanics

In statistical mechanics, we try to find average properties of a system by introducing appropriate distribution functions. If the system is completely isolated and if we make an a priori probability assumption that equal elementary volumes in phase space have equal probability, we may introduce the distribution function

$$f(\mathbf{p}, \mathbf{r}) = \text{const}, \qquad \mathscr{H}(\mathbf{p}, \mathbf{r}) = (E, E + \Delta E)$$
$$= 0, \qquad \mathscr{H}(\mathbf{p}, \mathbf{r}) \neq (E, E + \Delta E),$$

in view of the conservation law. Here, $\mathscr{H}(\mathbf{p}, \mathbf{r})$ is the Hamiltonian, $\mathbf{p}$ and $\mathbf{r}$ represent momenta and coordinates of the particles in the system, and $(E, E + \Delta E)$ represents a narrow range between $\mathscr{H} = E$ and $\mathscr{H} = E + \Delta E$.

The above distribution is called ***microcanonical****. In this distribution, no environmental effects act on the system: The system is completely isolated. However, actual macroscopic systems have some interactions with their environments. Without such interactions, it would not even be possible to measure their properties. Indeed, no parts of the universe are perfectly isolated macroscopic systems.*

What kinds of probability distribution laws hold for such realistic cases? If an appropriate distribution function is found, how can it give macroscopic quantities? How is statistical mechanics connected with such areas of macroscopic physics as thermodynamics? These are the subjects of this chapter.

2.1. Phase Space and the Liouville Theorem

We consider a system of N particles which moves in accordance with Hamilton's equations of motion:

$$\dot{\mathbf{r}}_i = \frac{\partial \mathscr{H}}{\partial \mathbf{p}_i}, \qquad \dot{\mathbf{p}}_i = -\frac{\partial \mathscr{H}}{\partial \mathbf{r}_i}, \qquad i = 1, 2, \ldots, N. \tag{1.1}$$

Here, $\mathbf{r}_i$ and $\mathbf{p}_i$ are the coordinates and the conjugate momenta of the ith particle and $\mathscr{H}$ is the Hamiltonian. The solution to Eq. (1.1) may be represented by the so-called *phase point* of coordinates $\mathbf{p} = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N)$ and $\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ in the $6N$-dimensional space known as *phase space*. As time elapses, the phase point moves around in phase space according to the equations of motion, but its trajectory does not intersect itself at any point, because of the uniqueness of the solution of the differential equations.

In general, the number of particles of a macroscopic system is very large, and the equations of motion are too complicated for rigorous solution. Moreover, in most cases, the precise initial conditions needed to solve the differential equations are not known. Furthermore, if we are concerned with macroscopic properties of the system, only the average behavior of the particles is important. Therefore, we try to describe the system in terms of a distribution function $f^{(N)}$. The probability of finding a phase point in the volume element $d\mathbf{p}^N\, d\mathbf{r}^N$ about the point $(\mathbf{p}, \mathbf{r})$ at time t is $f^{(N)}(\mathbf{p}, \mathbf{q}; t)\, d\mathbf{p}^N\, d\mathbf{r}^N$. Here, the vectors $\mathbf{p}$ and $\mathbf{r}$ stand for the set $\mathbf{p}_i$ and $\mathbf{r}_i$, $i = 1, 2, \ldots, N$, respectively, and

$$d\mathbf{p}^N\, d\mathbf{r}^N = \prod d\mathbf{p}_i\, d\mathbf{r}_i. \tag{1.2}$$

The distribution function $f^{(N)}$ defines an average of a given function $G(\mathbf{p}, \mathbf{r})$ in the phase space as follows:

$$\langle G \rangle = \frac{1}{\Omega} \int G(\mathbf{p}, \mathbf{r}) f^{(N)}(\mathbf{p}, \mathbf{r}; t)\, d\mathbf{p}^N\, d\mathbf{r}^N, \tag{1.3}$$

where

$$\Omega = \int f^{(N)}(\mathbf{p}, \mathbf{r}; t)\, d\mathbf{p}^N\, d\mathbf{r}^N \tag{1.4}$$

is the normalization constant of $f^{(N)}$. Note here that $\langle G \rangle$ may depend on time t.

If there is no sink or source, the phase points entering a volume must also leave in order to meet the conservation law. Thus,

$$\frac{\partial f^{(N)}}{\partial t} + \operatorname{div} \mathbf{j} = 0, \tag{1.5}$$

where $\mathbf{j}$ is the current vector defined by

$$\mathbf{j} = (f^{(N)}\dot{\mathbf{p}}_1, f^{(N)}\dot{\mathbf{p}}_2, \ldots, f^{(N)}\dot{\mathbf{r}}_N), \tag{1.6}$$

where

$$\dot{\mathbf{p}}_i = \dot{\mathbf{p}}_i(\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{r}_1, \mathbf{r}_2, \ldots; t); \qquad \dot{\mathbf{r}}_i = \dot{\mathbf{r}}_i(\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{r}_1, \mathbf{r}_2, \ldots; t).$$

Explicitly, Eq. (1.5) is

$$\frac{\partial f^{(N)}}{\partial t} + \sum_i \left(\frac{\partial}{\partial \mathbf{p}_i} \cdot f^{(N)} \dot{\mathbf{p}}_i + \frac{\partial}{\partial \mathbf{r}_i} \cdot f^{(N)} \dot{\mathbf{r}}_i \right) = 0,$$

or

$$\frac{\partial f^{(N)}}{\partial t} + \sum_i \left(\frac{\partial f^{(N)}}{\partial \mathbf{p}_i} \cdot \dot{\mathbf{p}}_i + \frac{\partial f^{(N)}}{\partial \mathbf{r}_i} \cdot \dot{\mathbf{r}}_i \right) + f^{(N)} \sum_i \left(\frac{\partial}{\partial \mathbf{p}_i} \cdot \dot{\mathbf{p}}_i + \frac{\partial}{\partial \mathbf{r}_i} \cdot \dot{\mathbf{r}}_i \right) = 0.$$

The last term is zero, from the equations of motion:

$$\frac{\partial}{\partial \mathbf{p}_i} \cdot \dot{\mathbf{p}}_i + \frac{\partial}{\partial \mathbf{r}_i} \cdot \dot{\mathbf{r}}_i = -\frac{\partial^2 \mathcal{H}}{\partial \mathbf{r}_i \, \partial \mathbf{p}_i} + \frac{\partial^2 \mathcal{H}}{\partial \mathbf{p}_i \, \partial \mathbf{r}_i} = 0.$$

Thus, we arrive at

$$\frac{\partial f^{(N)}}{\partial t} + [f^{(N)}, \mathcal{H}] = 0. \tag{1.7}$$

Here,

$$[f^{(N)}, \mathcal{H}] = \sum_i \left(\frac{\partial f^{(N)}}{\partial \mathbf{r}_i} \cdot \frac{\partial \mathcal{H}}{\partial \mathbf{p}_i} - \frac{\partial f^{(N)}}{\partial \mathbf{p}_i} \cdot \frac{\partial \mathcal{H}}{\partial \mathbf{r}_i} \right) \tag{1.8}$$

is the Poisson bracket. Equation (1.7) is equivalent to

$$\frac{df^{(N)}}{dt} = 0. \tag{1.9}$$

The conservation of the probability density expressed by Eq. (1.9) is called *Liouville's theorem*. The partial differential equation (1.7) determines, in principle, $f^{(N)}$. However, having to solve this equation is equivalent to having to solve the equations of motion (1.1), because its Lagrange subsidiary equations are just the equations of motion. Nevertheless, being a single equation for the distribution function, Eq. (1.9) is very useful for many purposes. For instance, the equation to determine the time change of $\langle G \rangle$ defined by Eq. (1.3) may be derived simply by multiplying Eq. (1.9) by G and integrating over the phase space.

Another important point is that any function of energy $\mathcal{H}(\mathbf{p}, \mathbf{q})$ alone satisfies the Liouville equation (1.7). The distribution function of a *microcanonical ensemble*, that is, a collection of similar systems having a definite energy, may be represented by

$$f(\mathcal{H}) = \delta(\mathcal{H} - E). \tag{1.10}$$

Using the Fourier integral representation of the Dirac δ-function, the average of a physical quantity G is given by

$$\langle G \rangle = \frac{1}{2\pi\Omega} \int_{-\infty}^{\infty} dx \, e^{ixE} \int G(\mathbf{p}, \mathbf{r}) e^{-ix\mathcal{H}} \, d\mathbf{p}^N \, d\mathbf{r}^N. \tag{1.11}$$

In this case, Ω is the total area of the energy surface.

The Liouville theorem may be expressed in the following forms. First, if a phase point that was at $(\mathbf{p}_0, \mathbf{r}_0)$ at time t is at a position $(\mathbf{p}, \mathbf{r})$ at time $t + s$, then

$$f^{(N)}(\mathbf{p}, \mathbf{r}; t + s) = f^{(N)}(\mathbf{p}_0, \mathbf{r}_0; t), \tag{1.12}$$

where, from the equations of motion, we have the relations

$$\mathbf{p} = \mathbf{p}(\mathbf{p}_0, \mathbf{r}_0; s), \qquad \mathbf{r} = \mathbf{r}(\mathbf{p}_0, \mathbf{r}_0; s). \tag{1.13}$$

Second, according to the Liouville theorem, or applying directly the equations of motion, we can prove that

$$\frac{\partial(\mathbf{p}, \mathbf{r})}{\partial(\mathbf{p}_0, \mathbf{r}_0)} = 1, \tag{1.14}$$

that is, the Jacobian for the transformation from $\mathbf{p}_0, \mathbf{r}_0$ to $\mathbf{p}, \mathbf{r}$ is unity, indicating the incompressibility of a volume element of phase space. This result can be understood by considering a small volume element about $\mathbf{p}_0, \mathbf{r}_0$ at time t. At time $t + s$, all the phase points in this volume will move to new positions and will form a new small volume about $\mathbf{p}, \mathbf{r}$. During the motion, there will be no crossing of trajectories. In particular, the trajectories formed by the surface points on the volume element are not crossed by any trajectory of a phase point irrespective of whether it is inside or outside. Therefore, the two volume elements must contain the same number of phase points. On the other hand, the density of phase points in the volume elements must be the same, according to Eq. (1.12). Thus, the volume itself remains constant. This leads to the above expression for the Jacobian. A special case of Eq. (1.14) has already been seen in Eq. (1.3) of Chapter 1.

2.2.* Ergodic Theories

The purpose of statistical mechanics is to evaluate macroscopic quantities such as thermodynamic functions from molecular dynamical observables. Therefore, let us examine a statistical definition of an experimentally observed quantity.

Any measurement of a physical quantity G takes a certain amount of time: It cannot be done instantaneously. We may consider that a time average

$$\langle G\rangle_\tau = \frac{1}{\tau}\int_{t_0}^{t_0+\tau} G(P_t)\,dt \tag{2.1}$$

is what is observed for a dynamical quantity $G(P_t)$, which is a function of a phase point P_t at time t. In this expression, the observation is made for the time interval between t_0 and $t_0 + \tau$. When the system is in equilibrium, we expect that $\langle G\rangle_\tau$ will be independent of the initial time t_0, and the averaging may be extended over an infinite time interval.

However, in practice, time averages may not be computable because the equations of motion are generally not soluble. The *ergodic theories* try to circumvent this difficulty by proving that time averages are replaceable by "ensemble averages."

The word "ergode" introduced by Boltzmann came from the Greek words ἔργον (work), from which the word energy has been formed, and ὁδός (path). Boltzmann did not use "ergode" as it is used now, but rather for a path on an energy surface, although he did vaguely use the concept of what we call the "quasi-ergodic hypothesis."

If the trajectory of a dynamical motion covers the entire energy surface uniformly (*ergodic hypothesis*), the time average $\langle G\rangle_\tau$ may be replaced by an ensemble average of the type

$$\langle G\rangle_\tau = \langle G\rangle = \frac{\int G(P)\,d\Gamma}{\int d\Gamma}, \tag{2.2}$$

where $d\Gamma$ is the volume element of the phase space, and the integration is carried over all the points between the energy surfaces $\mathcal{H} = E$ and $\mathcal{H} = E + \Delta E$.

The ergodic hypothesis caught the attention of mathematicians, and a new, important branch of mathematics called ergodic theory grew along with the development of measure theory. The main problem in ergodic theory was to find whether or not time averages existed, and if so, whether or not they could be replaced by ensemble averages. The mere existence of time averages, which is assumed in physics, is not an obvious fact in mathematics, and had to be studied.

It was soon found that the ergodic hypothesis was not valid. From a topological observation, it is clear that no single trajectory can fill out the whole energy surface. Thus, attempts have been made to establish the *quasi-ergodic hypothesis*, which states that a single phase trajectory, although it does not completely fill out the energy surface, comes arbitrarily close to a given point on the surface. If in most physical systems a trajectory in the phase space covers the energy surface densely enough, then the replacement of averages might still be possible.

After many years of confusion and obscurity, two major contributions were made, almost at the same time, by von Neumann (*1*) and Birkoff (*2*). In 1932, von Neumann proved the so-called *mean ergodic theorem* and Birkoff, in 1931, proved the *individual ergodic theorem* (*3*, *4*). According to the latter, the limiting time average

$$\langle G\rangle_\infty = \lim_{\tau\to\infty}\frac{1}{\tau}\int_0^\tau G(P_t)\,dt \tag{2.3}$$

exists for all points P of an *invariant part* (to be discussed later) of the phase space except for at most a certain set of measure zero. He then found the condition under which $\langle G\rangle_\infty$ is replaceable by $\langle G\rangle$. On the other hand, von Neumann's theorem proves a convergence to a limit $\langle G\rangle_\infty$ such that

$$\lim_{\tau\to\infty} |\langle G\rangle_\tau - \langle G\rangle_\infty|^2 \to 0.$$

An invariant part of phase space means that an arbitrary point in this part remains within that part during all the natural motion. When a phase space cannot be decomposed into two invariant parts, we say that it is *metrically indecomposable*. Birkoff proved the individual ergodic theorem under the necessary and sufficient condition that the phase space is metrically indecomposable. His theorem states that almost everywhere in the phase space (namely, except for a part of measure zero)

$$\langle G\rangle_\infty = \langle G\rangle. \tag{2.4}$$

The outline of Birkoff's proof is as follows. He first shows that, almost everywhere, the time average $\langle G\rangle_\infty$ exists and is constant:

$$\langle G\rangle_\infty = \text{const.} \tag{2.5}$$

If the time average were not constant, one could decompose the phase space into two parts such that $\langle G\rangle_\infty > \alpha$ in V_1 and $\langle G\rangle_\infty \leqslant \alpha$ in V_2. Such a decomposition is then found to contradict the assumption of metrical indecomposability. The next step is to find whether $\langle G\rangle_\infty$ is replaceable by $\langle G\rangle$. It is difficult to see this without imposing the appropriate mathematical rigor, but

very roughly, the proof goes as follows:

$$\langle G\rangle_\tau = \frac{\int \langle G\rangle_\tau \, d\Gamma}{\int d\Gamma} = \frac{\int [\langle G\rangle_\tau - \langle G\rangle_\infty] \, d\Gamma}{\int d\Gamma} + \frac{\int \langle G\rangle_\infty \, d\Gamma}{\int d\Gamma}. \tag{2.6}$$

In the limit $\tau \to \infty$, the first term may be considered to vanish almost everywhere. In the second term, the limiting process and the integration may be interchanged. Using the Liouville theorem $d\Gamma = d\Gamma_t$, we then find

$$\frac{\int \langle G\rangle_\tau d\Gamma}{\int d\Gamma} = \frac{1}{\tau} \frac{\int G(P_t)\, dt\, d\Gamma_t}{\int d\Gamma_t}$$

$$= \frac{1}{\tau} \int \langle G\rangle \, dt = \langle G\rangle.$$

The Birkoff theorem requires the condition of metrical indecomposability. The question of whether or not this condition is satisfied in most physical systems is difficult to answer, but it seems that at least a system of hard spheres might satisfy this requirement (*5*).

A difficulty arises in this respect if there is another integral of the equations of motion. Suppose $\phi = \phi(P)$ is such an integral. This integral cannot be constant over all the energy surface, because by assumption ϕ is an independent integral. If $\phi > \alpha$ in V_1 and $\phi \leqslant \alpha$ in V_2, then V_1 and V_2 may be invariant parts of the phase space, since ϕ is an integral. Thus, the surface of a constant energy may not be metrically indecomposable.

One can circumvent this difficulty by various considerations. First, the additional integral may not give rise to domains of finite volume: The dimension of the hypersurface $\phi = \phi(P)$ could be much less than that of the energy surface. Second, physically meaningful phase functions might be more complicated than mathematical phase functions. The ergodic theorem is for arbitrary (Lebesgue-measurable) phase functions, but the phase functions entering in physics satisfy many conditions. For instance, they must assume the same value at different phase points that correspond to the same macroscopic physical state. If the energy surface cannot be split into invariant subspaces such that points of the same physical state belong to the same part, the ergodic theorem might still be valid. Furthermore, the physical ergodicity might be meaningful in the limit when the number of particles approaches infinity.

Some comments are now due concerning the ensemble average. So long as the energy is kept constant, it is appropriate to consider a small part of the entire phase space which lies between the two energy surfaces $\mathscr{H} = E$ and $\mathscr{H} = E + \Delta E$, where ΔE is a small energy increment. This region is invariant and has the volume element $dS\, d\mathscr{H}/|\operatorname{grad} \mathscr{H}|$, where dS is the surface element and $d\mathscr{H}/|\operatorname{grad} \mathscr{H}|$ is its normal. In the limit, when the width of the

layer becomes infinitely small, we obtain the surface element

$$d\sigma = \frac{1}{\Delta E} \left[d\mathscr{H} \frac{dS}{|\operatorname{grad} \mathscr{H}|} \right]_{\mathscr{H}=E}$$

$$= \left[\frac{dS}{|\operatorname{grad} \mathscr{H}|} \right]_{\mathscr{H}=E}. \tag{2.7}$$

In terms of $d\sigma$, a microcanonical ensemble average may be defined as

$$\langle G\rangle = \frac{\int_{\mathscr{H}=E} G(P)\, d\sigma}{\int d\sigma}, \tag{2.8}$$

where

$$d\sigma = \frac{dS}{|\operatorname{grad} \mathscr{H}|}. \tag{2.9}$$

The motion of a phase point forms a transformation group: If we denote the motion from P to P_t by T_t, then

$$T_{t_1+t_2} = T_{t_1} T_{t_2}, \qquad T_0 = 1, \qquad T_t T_t^{-1} = 1. \tag{2.10}$$

Let us consider a phase function $G(P)$ of a phase point P which moves on the energy surface in accordance with T_t. We introduce the following notations:

$$G(P_t) = T_t G(P),$$

$$(G, F) = \int GF^* \, d\sigma, \tag{2.11}$$

$$(G, G) = \int GG^* \, d\sigma = |G|^2,$$

where $G = G(P)$, and $F = F(P)$ is another phase function. If

$$(G(P_t), G(P_t)) = (G(P), G(P)) = |G|^2 = \text{const}, \tag{2.12}$$

that is, if the transformation T_t does not change the absolute magnitude of G, then $T_t T_t^* = 1$, because of the Liouville theorem $d\sigma_t = d\sigma$. Namely, the transformations form a unitary group.

For such transformations, let us examine the time average

$$\frac{1}{\tau} \int_0^\tau G(P_t)\, dt = \frac{1}{\tau} \int_0^\tau T_t G(P)\, dt. \tag{2.13}$$

The period τ for the average may conveniently be chosen to be an integer.

Because of the additivity of T_t, we have

$$\frac{1}{\tau}\int_0^\tau T_t\,dt = \frac{1}{\tau}\int_0^1 (1 + T_1 + T_1^2 + \cdots + T_1^{\tau-1})T_t\,dt$$

$$= \frac{1}{\tau}\int_0^1 \frac{T_1^\tau - 1}{T_1 - 1}\,T_t\,dt. \tag{2.14}$$

Consider now that G is split into two parts such that

$$G = G_1 + G_2, \qquad (G_1, G_2) = 0. \tag{2.15}$$

Let us assume that G_1 is given by another phase function as follows:

$$G_1 = T_1F - F. \tag{2.16}$$

Then,

$$\frac{1}{\tau}\int_0^\tau T_t G_1(P)\,dt = \frac{1}{\tau}\int_0^1 (T_1^\tau - 1)T_t F\,dt.$$

In the notation of Eq. (2.12), we find

$$\left|\frac{1}{\tau}\int_0^\tau T_t G_1\,dt\right|^2 \le \frac{1}{\tau}\left[\left|\int_0^1 T_1^\tau T_t F\,dt\right|^2 + \left|\int_0^1 T_t F\,dt\right|^2\right]$$

$$= \frac{1}{\tau}\,2\left|\int_0^1 T_t F\,dt\right|^2.$$

The last equality is due to the unitary of T_1. Therefore, for $\tau \to \infty$, the time average of G_1 vanishes:

$$\lim_{\tau\to\infty} \langle G_1\rangle_\tau = 0.$$

Using the unitarity of T_1 and introducing Eq. (2.16) into Eq. (2.15), we find

$$(T_1F - F, G_2) = (F, T_1^{-1}G_2) - (F, G_2) = (F, T_1^{-1}G_2 - G_2) = 0.$$

If this equality holds for all F except for the case $(F, F) = 0$, then

$$T_1G_2 = G_2.$$

That is, G_2 is invariant under the operation T_1.

Therefore, we find from Eq. (2.14)

$$\frac{1}{\tau}\int_0^\tau T_t G_2\,dt = \frac{1}{\tau}\int_0^1 \tau T_t G_2\,dt = \int_0^1 T_t G_2\,dt.$$

The right-hand side of this equation gives the time average

$$\langle G\rangle_\infty = \lim_{\tau\to\infty}\frac{1}{\tau}\int_0^\tau T_t G\,dt$$

$$= \int_0^1 T_t G_2\,dt \tag{2.17}$$

This time average is invariant under the transformation T_t. Since T_t operates on the phase-space function, we find

$$T_t\langle G\rangle_\tau \equiv \frac{1}{\tau}\int_0^\tau T_{t+t_1}G\,dt_1$$

$$= \frac{t+\tau}{\tau}\frac{1}{t+\tau}\int_0^{t+\tau} T_{t+t_1}G\,d(t+t_1) - \frac{1}{\tau}\int_\tau^{t+\tau} T_{t+t_1}G\,dt_1. \tag{2.18}$$

In the limit $\tau \to \infty$, we have

$$T_t\langle G\rangle_\infty = \langle G\rangle_\infty. \tag{2.19}$$

The invariance of $\langle G\rangle_\infty$ leads to

$$(G, 1) = (\langle G\rangle_\infty, 1), \tag{2.20}$$

from the Liouville theorem. Using the invariance property of $\langle G_2\rangle_\tau$, we can write this equation as follows:

$$\langle G\rangle_\infty = \frac{\int G\,d\sigma}{\int d\sigma}, \tag{2.21}$$

where the right-hand side integrals are over the surface of constant energy. Summarizing, we have observed that if a function G_2 is an invariant function in phase space, $\langle G_2\rangle_\infty$ exists and is equal to $\langle G_2\rangle$. This conclusion is based on a *mean* convergence such as expressed in Eq. (2.20).

Example. We consider the ergodicity of a rectilinear motion in a two-dimensional rectangle with periodic conditions.

Without loss of generality, the rectilinear motion may be assumed to start from the origin:

$$x(t) = t, \qquad y(t) = \omega t; \tag{2.22}$$

x and y take on any values in the domain with periodicity 1:

$$\begin{aligned} x + 1 &\equiv x \pmod 1,\\ y + 1 &\equiv y \pmod 1. \end{aligned} \tag{2.23}$$

If ω is a rational number, one can choose integers a and b such that

$$\omega = \frac{b}{a}. \tag{2.24}$$

At $t = a$, $x = a \equiv 0$ and $y = \omega a = b \equiv 0$. Thus, at this time, the point returns to the origin.

If ω is irrational, one can prove that the trajectory can come to a close vicinity of any given point in the domain. This property, called quasi ergodicity, may be stated mathematically as follows: For an arbitrarily chosen point (x_0, y_0), the distances

$$x - x_0 = t - x_0, \qquad y - y_0 = \omega t - y_0 \tag{2.25}$$

are z ro with modulus 1.

For our purpose, we introduce an analytic function $f(x, y)$. In view of the periodicity, we expand the function in a Fourier series:

$$f(x, y) = \sum_{m, n} a_{mn} e^{2\pi i(mx+ny)}.$$

Then, we observe that

$$\frac{1}{t}\int_0^t f(x, y)\, dt = \sum_{m,\, n\neq 0} a_{mn} \frac{1}{t} \frac{e^{2\pi i(m+n\omega)t} - 1}{2\pi i(m + n\omega)} + a_{00}. \tag{2.26}$$

Since $m + n\omega \neq 0$ except for $m = n = 0$, the factor of a_{mn} vanishes in the limit $t \to \infty$. Thus,

$$\lim_{t\to\infty} \frac{1}{t}\int_0^t f(x(t), y(t))\, dt = a_{00}; \tag{2.27}$$

i.e.,

$$\langle f(x(t), y(t))\rangle_\infty = \int_0^1 \int_0^1 f(x, y)\, dx\, dy.$$

Use of this result for $f(x(t), y(t))$ of Eq. (2.25) will demonstrate the quasi ergodicity.

2.3. H-Theorem for Systems in Equilibrium

Since Boltzmann's presentation in 1872 of the famous H-theorem in order to discuss how equilibrium is approached, it has encountered many objections and engendered much discussion and modification. We now have roughly the following four versions of the H-theorem, all of which try to show the entropy increase toward statistical equilibrium, but differing from each other in content (*6*):

(*a*) *Dynamical approach.*
- (1) H-theorem based on Boltzmann's equation.
- (2) H-theorem based on the reduced Liouville equations and the master equation.

(*b*) *Statistical approach.*
- (3) H-theorem concerned with the canonical distribution.
- (4) H-theorem due to coarse-graining or information loss.

Using the H-theorem, Boltzmann showed that the detailed balance condition was necessary for equilibrium. The Maxwell–Boltzmann distribution function follows from this condition. Since the Maxwell–Boltzmann distribution satisfies the Boltzmann equation, Boltzmann proved, in effect, that the distribution was the necessary and sufficient expression for equilibrium.

Boltzmann's H-theorem is proved by the simple exchange of the two velocity integral variables necessary to express $-d\langle \log f\rangle/dt$ in terms of the Boltzmann collision integral.

The irreversibility in Boltzmann's H-theorem led to many questions and objections because it appeared to contradict dynamics, which is reversible in time. To answer such questions, Boltzmann and later Ehrenfest (*8*) gave the interpretation that H decreases statistically rather than dynamically.

Boltzmann's H-theorem played a historically important role for the development of statistical mechanics. Concerning its content, it is important to recognize that the Boltzmann H-theorem determines an equilibrium expression only for a one-body distribution function. It is clear that an N-body system may not adequately be described by such a one-body distribution function. Therefore, we shall discuss the H-theorem due to Gibbs, which determines the equilibrium form of an N-body distribution function.

Let us introduce

$$H = \frac{1}{\Omega}\int f^{(N)} \log f^{(N)}\, d\mathbf{r}^N\, d\mathbf{p}^N,$$
$$\Omega = \int f^{(N)}\, d\mathbf{r}^N\, d\mathbf{p}^N. \tag{3.1}$$

From Eq. (3.1), H is equal to the average of $\log f$:

$$H = \langle \log f\rangle. \tag{3.2}$$

We notice that the H-function is not a function of phase-space coordinates $(\mathbf{r}, \mathbf{p})$, since the phase-space integration has already been carried out in Eq. (3.1). However, through the coordinate-space integrations, H may depend on

the total volume of the system. It may also be a function of conditions externally imposed on the system.

According to the Liouville theorem, the H-function should not change with time. However, the value of H depends on the function $f^{(N)}$. If $f^{(N)}$ is restricted by an energy condition, the following statistical H-theorem holds.

H-THEOREM. *If the distribution function satisfies the condition*

$$\frac{1}{\Omega}\int \mathscr{H} f^{(N)}\, d\mathbf{r}^N\, d\mathbf{p}^N = U, \tag{3.3}$$

$\mathscr{H}$ is a minimum when

$$f^{(N)} = e^{\beta(F-\mathscr{H})}. \tag{3.4}$$

Here, U, F, and β are constants independent of phase-space coordinates.

Proof. We simplify the notation by writing $f^{(N)}$ as f. Our procedure of proof is to compare H-functions for $f_{\mathrm{I}} = e^{\beta(F-\mathscr{H})}$ and for an arbitrary function f_{II} with the same normalization. Introducing g by writing $f_{\mathrm{II}} = f_{\mathrm{I}} g$, we have

$$H_{\mathrm{II}} - H_{\mathrm{I}} = \frac{1}{\Omega}\int [f_{\mathrm{II}}(\log f_{\mathrm{II}}) - f_{\mathrm{I}} \log f_{\mathrm{I}}]\, d\mathbf{r}^N\, d\mathbf{p}^N$$

$$= \frac{1}{\Omega}\int f_{\mathrm{I}}[g(\log f_{\mathrm{I}} g) - \log f_{\mathrm{I}}]\, d\mathbf{r}^N\, d\mathbf{p}^N, \tag{3.5}$$

because of the normalization condition

$$\int (f_{\mathrm{II}} - f_{\mathrm{I}})\, d\mathbf{r}^N\, d\mathbf{p}^N = \int f_{\mathrm{I}}(g-1)\, d\mathbf{r}^N\, d\mathbf{p}^N = 0. \tag{3.6}$$

With the condition of Eq. (3.3),

$$\frac{1}{\Omega}\int (f_{\mathrm{II}} - f_{\mathrm{I}})\mathscr{H}\, d\mathbf{r}^N\, d\mathbf{p}^N = 0, \tag{3.7}$$

and with $\beta(F - \mathscr{H}) = \log f_{\mathrm{I}}$, Eqs. (3.6) and (3.7) give

$$\frac{1}{\Omega}\int (f_{\mathrm{II}} - f_{\mathrm{I}}) \log f_{\mathrm{I}}\, d\mathbf{r}^N\, d\mathbf{p}^N = \frac{1}{\Omega}\int f_{\mathrm{I}}(g-1)\log f_{\mathrm{I}}\, d\mathbf{r}^N\, d\mathbf{p}^N$$

$$= 0. \tag{3.8}$$

Subtracting Eqs. (3.6) and (3.8) from Eq. (3.5), we obtain

$$H_{\mathrm{II}} - H_{\mathrm{I}} = \frac{1}{\Omega}\int f_{\mathrm{I}}[g(\log g) - g + 1]\, d\mathbf{r}^N\, d\mathbf{p}^N. \tag{3.9}$$

From its definition, f_{I} is positive, and the quantity $[g(\log g) - g + 1] = \int_1^g \log g\, dg$ cannot be negative for $g > 0$ irrespective of the magnitude of g (Fig 2.1). Thus, H_{II} is never smaller than H_{I}.

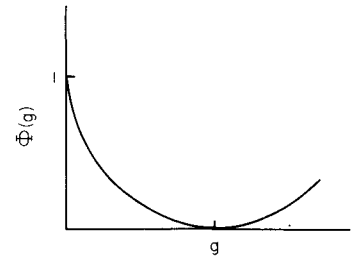


Fig. 2.1. The nonnegative function $\Phi(g) = g(\log g) - g + 1$.

When H is a minimum, we consider that the system is in equilibrium. A decreasing H corresponds to approaching equilibrium. Since the thermodynamic entropy increases toward equilibrium, we adopt

$$S = -kH, \tag{3.10}$$

that is, the thermodynamic entropy is proportional to the H-function.

The proportionality constant k must be a universal constant. Suppose that a system consists of two independent subsystems 1 and 2. The total distribution function must be the product of the distribution functions of the subsystems:

$$f(x, y) = f_1(x) f_2(y),$$

where x and y are representative phase-space variables. By definition,

$$\langle \log f \rangle = \langle \log f_1 \rangle_1 + \langle \log f_2 \rangle_2 .$$

The two terms on the right-hand side are independent of each other. Therefore, if each term is related to the entropy, the proportionality constant must be independent of the systems. The constant k is Boltzmann's constant.

The distribution function given by Eq. (3.4) is called the *canonical distribution function*. It corresponds to thermal equilibrium. In the next chapter, we shall derive the same distribution function by a different method.

2.4. Meanings of the Constants in the Canonical Distribution Function

The canonical distribution function includes two constants F and β which are independent of phase-space coordinates. Let us find the meanings of these constants. Using Eqs. (3.3) and (3.4), we find

$$U = \frac{\int \mathscr{H} e^{-\beta\mathscr{H}}\, d\mathbf{r}^N\, d\mathbf{p}^N}{\int e^{-\beta\mathscr{H}}\, d\mathbf{r}^N\, d\mathbf{p}^N}. \tag{4.1}$$

In other words, U is the average energy. If the Hamiltonian does not include the kinetic energy due to the mass motion of the system, U may be adopted as the *internal energy*. If the Hamiltonian does not include external forces, U is a function of β and V:

$$U = U(\beta, V).$$

Equation (4.1) can be applied to an ideal gas. The Hamiltonian is independent of coordinates, and the coordinate integrations in Eq. (4.1) yield V^N. We arrive at

$$U = \frac{3N}{2\beta}.$$

On the other hand, the absolute temperature scale is defined such that $U = 3NkT/2$. Therefore, we conclude that

$$\beta = \frac{1}{kT}. \tag{4.2}$$

This identification of β is meaningful because Eq. (4.1) gives

$$\left[\frac{\partial U}{\partial(-\beta)}\right]_V = \langle \mathscr{H}^2 \rangle - \langle \mathscr{H} \rangle^2 = \langle (H - \langle H \rangle)^2 \rangle \tag{4.3}$$

which is in general positive. That is, U increases with $-\beta$. Suppose now there are two systems with the same Hamiltonian and volume but with different β's, β_1 and β_2. Then,

$$U(\beta_1) \neq U(\beta_2)$$

if $\beta_1 \neq \beta_2$. Therefore, β is a parameter that can be used to measure the internal energy.

On the other hand, from Eqs. (3.1) and (3.4), we find

$$e^{-\beta F} = \frac{1}{\Omega} \int e^{-\beta \mathscr{H}} \, d\mathbf{r}^N \, d\mathbf{p}^N. \tag{4.4}$$

F is uniquely determined if the normalization constant Ω is given. The constant Ω can be 1, or, as we shall see later, it can be $N!$ for an N particle system. [see Eq. (1.4) of Chapter 6]. Then,

$$F = F(\beta, V). \tag{4.5}$$

From Eq. (3.4), we obtain

$$\beta(F - \mathscr{H}) = \log f^{(N)}. \tag{4.6}$$

Multiplying both sides by $f^{(N)}/\Omega$ and integrating over the entire phase space, we find

$$\begin{aligned} F &= U + \frac{H}{\beta} \\ &= U - \frac{S}{k\beta}. \end{aligned} \tag{4.7}$$

In thermodynamics, we call

$$F = U - TS \tag{4.8}$$

the Helmholtz free energy. Taking U, S, and β as the internal energy, the entropy, and the reciprocal thermal energy, we identify F given by Eq. (4.4) as the Helmholtz free energy.

We can understand the identification of F of Eq. (4.4) as the Helmholtz free energy in a slightly different way. Let us take the total derivative of F of Eq. (4.7). Since F is defined by a phase-space integral as in Eq. (4.4), it must be a function of the total volume V and β. In addition, if the Hamiltonian includes external fields, F depends also on the fields. For a fixed total number of particles, the total derivative of F is

$$dF = -\frac{d\beta}{\beta}(F - U) + dU - T\,dS. \tag{4.9}$$

If U is the internal energy, the first and second laws of thermodynamics give

$$dU = T\,dS - p\,dV. \tag{4.10}$$

Thus, Eq. (4.9) becomes

$$dF = -\frac{d\beta}{\beta}(F - U) - p\,dV. \tag{4.11}$$

Adopting the temperature scale given by Eq. (4.2), we find that Eq. (4.11) has the same form as the total derivative of the thermodynamically defined Helmholtz free energy:

$$dF = -S\,dT - p\,dV. \tag{4.12}$$

Therefore, F of Eq. (4.4) gives the statistical free energy.

In thermodynamics, several thermodynamic functions are used depending on independent variables. For instance, the Gibbs free energy is convenient when the temperature T and pressure p are chosen as independent variables.

According to our definition, F is a function of β and V. The thermodynamic function that has temperature and volume as independent variables is the Helmholtz free energy.

The absolute temperature scale has been introduced in such a way that the equation of state of an ideal gas is given by

$$pV = NkT. \tag{4.13}$$

Let us confirm the role of F by applying directly Eq. (4.4) to an ideal gas. We find

$$\begin{aligned} e^{-\beta F} &= V^N \frac{1}{\Omega}\left[\int \exp\left(-\frac{\beta p^2}{2m}\right) d\mathbf{p}\right]^N \\ &= \frac{V^N}{\Omega}\left(\frac{2m\pi}{\beta}\right)^{3N/2} \end{aligned} \tag{4.14}$$

Thus, $F \approx -(N/\beta)\log V$ and since $p = -(\partial F/\partial V)_T$ according to Eq. (4.11), we obtain

$$pV = \frac{N}{\beta}.$$

This confirms that $\beta = 1/kT$ and that F is the free energy.

2.5. Coarse-Graining

In Section 2.3, the H-function for the canonical distribution was compared with that for an arbitrary distribution function and was found to be smaller than or equal to the latter. From this H-theorem, we concluded that the canonical distribution is the equilibrium distribution. However, in the discussion, no time change was considered. Therefore, let us try to find whether the entropy increases toward equilibrium or not. For this purpose, we must introduce a special mechanism to consider a time change of the entropy, because the Liouville theorem requires that the H-function evaluated for the total distribution function of a system does not change with time. For clarity, let us start with a simple mathematical inequality which can be understood quite independent of our next step—the discussion of entropy increase with time.

The inequality is stated as follows:

THEOREM. *For two distribution functions $f(x)$ and $\hat{f}(x)$ in Γ-space satisfying the conditions*

$$\int_\Gamma f(x)\,dx = \int_\Gamma \hat{f}(x)\,dx; \qquad f, \hat{f} \geqslant 0, \tag{5.1}$$

define

$$\begin{aligned} H &= \int_\Gamma f \log f\, dx, \\ \hat{H} &= \int_\Gamma f \log \hat{f}\, dx. \end{aligned} \tag{5.2}$$

Then, the following inequality holds (7):

$$\int_\Gamma f \log f\, dx \geqslant \int_\Gamma f \log \hat{f}\, dx. \tag{5.3}$$

We shall call this the ***Gibbs inequality****.*

Proof. The proof of this inequality may be given following the Gibbs method discussed in the previous section. Using Eq. (5.1) and introducing a new function g by writing $f = \hat{f}g$, we obtain

$$\begin{aligned} H - \hat{H} &= \int_\Gamma [f(\log f - \log \hat{f}) - (f - \hat{f})]\, dx \\ &= \int_\Gamma \hat{f}[g(\log g) - g + 1]\, dx. \end{aligned}$$

By definition and as is illustrated in Fig. 2.1, we find for $\hat{f} \neq 0$

$$g = \frac{f}{\hat{f}} \geqslant 0;$$

$$g(\log g) - g + 1 \geqslant 0.$$

Therefore,

$$H \geqslant \hat{H},$$

where the equality holds when $f = 0$ or

$$f = \hat{f}.$$

Note that the Gibbs inequality has nothing to do with a time change. It holds for any nonnegative functions with the same normalization. It is certainly not

true without the condition (5.1). However, if a time change has to be introduced, it is $\hat{H}$, which can change. Let us now introduce a special mechanism by which $\hat{H}$ changes with time (*8*).

A coarse-grained distribution function is defined by averaging a fine-grained distribution function $f(x)$ for small cells in phase space (*6*):

$$\hat{f}_i = \frac{\int_i f(x)\,dx}{\Delta x_i}. \tag{5.4}$$

Note that

$$\sum_i \hat{f}_i \Delta x_i = \sum_i \int_i f(x)\,dx = \int_\Gamma f(x)\,dx. \tag{5.5}$$

We introduce the $\overline{H}$-function defined in terms of the coarse-grained distribution function by

$$\overline{H} = \sum_i \hat{f}_i \Delta x_i \log \hat{f}_i. \tag{5.6}$$

Introducing $\hat{f}(x) = \hat{f}_i$ for x belonging to cell i, we find

$$\begin{aligned}\overline{H} &= \sum_i \left[\int_i f(x)\,dx\right] \log \hat{f}_i \\ &= \int_\Gamma f(x) \log \hat{f}(x)\,dx \\ &= \hat{H}.\end{aligned} \tag{5.7}$$

From the Gibbs inequality (5.3), we have

$$H \geqslant \overline{H}. \tag{5.8}$$

Thus, $\overline{H}$ is less than H due to coarse-graining except for $f = \hat{f}$.

In terms of the inequality (5.8), let us now try to prove that the $\overline{H}$-function decreases with time. We consider that the initial situation is described by the distribution function f which is uniform over each cell in the phase space. In other words, we choose

$$H(0) = \overline{H}(0) \qquad \text{at} \quad t = 0. \tag{5.9}$$

After some time, $t > 0$, $\hat{f}$ will spread in an intricate way over many cells so that $f \neq \hat{f}$ in general, and

$$H(t) > \overline{H}(t). \tag{5.10}$$

However, from the Liouville theorem,

$$H(t) = H(0) = \overline{H}(0). \tag{5.11}$$

Thus, for $t > 0$,

$$\overline{H}(0) > \overline{H}(t). \tag{5.12}$$

In other words, we have proven that $\overline{H}$ has decreased.

It is important to recognize that the H-theorem and, actually, all our statistical considerations are based on the *a priori* probability assumption that the phase points in the same phase-space volume dx about x have the same probability $f(x)$. Suppose, then, we divide the entire phase space into small cells in which the equal *a priori* probability assumption holds. As is illustrated in Fig. 2.2, the system at $t = 0$ is represented by the phase points in the cell i

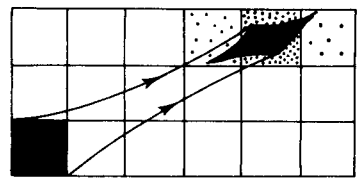


Fig. 2.2. Motion of a phase cell. Its shape changes but its area remains the same. The average densities of phase points in various cells in the new location are shown by dots.

(dark square), and the distribution function is not zero in this cell. At a later time t, every point in the cell i moves to a new point so that a new domain will be occupied. The shape of this new domain may be different from the original one, although the volume remains the same due to the Liouville theorem. Generally, the new domain spreads over many cells, and the average of $\log \hat{f}$ evaluated over these cells will be less than $f_i \Delta_i \log f_i$. If $f_i\,(< 1)$ decreases, corresponding to the spreading of the cell i, then $\log \hat{f}_j$ decreases on the average.

As a result of coarse-graining, a certain amount of information about a system becomes unavailable. Taking $-k\overline{H}$ as the entropy, the above result corresponds to the situation where the entropy increases due to information loss. One can obtain the coarse-grained distribution function $\hat{f}$ from the fine-grained distribution function f, but the reverse process is not possible. That is, the coarse-graining process introduces irreversibility.

Suppose that at an initial time we had a complete description of a system with the fine-grained distribution function f. At a later time, when we perform a new experiment on the system, we consider that only a coarse-grained distribution can be obtained because observation is limited. Then, the entropy evaluated by the coarse-grained distribution function is usually larger than that given by the one at the initial time.

The question may be asked: Is there a possibility of having the reverse situation? The answer is in the negative. Such a possibility is excluded so long as we cannot perform an experiment directed toward the past.

The H-theorem given in terms of the coarse-grained distribution function differs from the one we gave before. It does not specify the canonical distribution function for equilibrium. In this sense, the H-theorem in the form due to Gibbs is important. Nevertheless, we shall later see that our mathematical inequality has some other applications in statistical physics.

Problem

For the two distribution functions

$$f(x) = 1, \qquad \hat{f}(x) = 2(1 - x), \qquad 0 \leq x \leq 1,$$

which have the same normalization condition, evaluate

$$H = \int_0^1 f \log f \, dx,$$

$$\overline{H} = \int_0^1 \hat{f} \log \hat{f} \, dx,$$

$$\hat{H} = \int_0^1 f \log \hat{f} \, dx,$$

and compare the results with each other.

2.6. Product Approximation for the Distribution Function

Let us consider an N-particle system of the total distribution function $f^{(N)}(x, y)$. The phase-space variables x and y together specify completely the configurations of the system. For our purpose, x is considered to represent all the variables of the subset of n particles out of N and y is for the remaining $N - n$ particles. These subsets are described by the reduced distribution functions $f^{(n)}(x)$ and $f^{(N-n)}(y)$, respectively, which are obtained from $f^{(N)}(x, y)$ as follows:

$$\begin{aligned} \int f^{(N)}(x, y) \, dy &= f^{(n)}(x); \\ \int f^{(N)}(x, y) \, dx &= f^{(N-n)}(y). \end{aligned} \tag{6.1}$$

For convenience, we shall use the normalization

$$\int f^{(N)}(x, y) \, dx \, dy = 1. \tag{6.2}$$

Correspondingly, we have

$$\int f^{(n)}(x) \, dx = \int f^{(N-n)}(y) \, dy = 1. \tag{6.3}$$

We remark that our subdivision of the phase-space variables is not a special one. Although we are using the different notations x and y for the subsets, both sets of variables may represent physical quantities of the same character such as coordinates and momenta.

As is frequently done in treating many-body systems, we shall now approximate the total distribution function $f^{(N)}(x, y)$ by the product

$$f^{(n)}(x) f^{(N-n)}(y)$$

satisfying the normalization condition

$$\int f^{(n)}(x) f^{(N-n)}(y) \, dx \, dy = 1. \tag{6.4}$$

In terms of the correct and the approximate distribution functions, we define H-functions as follows:

$$\begin{aligned} H &= \int f^{(N)}(x, y) \log f^{(N)}(x, y) \, dx \, dy \\ H_a &= \int f^{(n)}(x) f^{(N-n)}(y) \log[f^{(n)}(x) f^{(N-n)}(y)] \, dx \, dy. \end{aligned} \tag{6.5}$$

Using Eqs. (6.1) and (6.3) in H_a, we find

$$\begin{aligned} H_a &= \int f^{(n)}(x) \log f^{(n)}(x) \, dx + \int f^{(N-n)}(y) \log f^{(N-n)}(y) \, dy \\ &= \int f^{(N)}(x, y) \log[f^{(n)}(x) f^{(N-n)}(y)] \, dx \, dy \end{aligned} \tag{6.6}$$

We note that both $f^{(N)}(x, y)$ and $f^{(n)}(x) f^{(N-n)}(y)$ are positive and satisfy the same normalization condition of the type (6.2). Thus, we can use the Gibbs inequality (5.3). We observe that

$$\begin{aligned} \Delta H &= H - H_a \\ &= \int f^{(N)}(x, y) \log f^{(N)}(x, y) \, dx \, dy \\ &\quad - \int f^{(n)}(x) \log f^{(n)}(x) \, dx - \int f^{(N-n)}(y) \log f^{(N-n)}(y) \, dy \\ &= \int f^{(N)}(x, y) \log f^{(N)}(x, y) \, dx \, dy \\ &\quad - \int f^{(N)}(x, y) \log[f^{(n)}(x) f^{(N-n)}(y)] \, dx \, dy, \end{aligned} \tag{6.7}$$

and therefore

$$H \geqslant H_a. \tag{6.8}$$

We have thus proven that the product assumption causes, in general, a decrease in the H-function. The proof is rigorous and is quite general in the sense that the character of the subdivision is not specified. Indeed, there could be many reduced distribution functions of arbitrary order. Whenever these reduced distribution functions are used to approximate the correct distribution function, some information on the system will be lost. This information loss appears as an entropy increase.

2.7. H-Theorem Based on the Master Equation

Let us consider a system which at a time t has the probability $P_i(t)$ of being in a state i. At a later time $t + \Delta t$, where $\Delta t > 0$ is small, the probability changes to $P_i(t + \Delta t)$. The time rate of change P_i may be quite generally determined by the so-called *master equation*, which is also called the *Pauli equation*.

The master equation is based on the assumption

$$P_i(t + \Delta t) = \sum_k A_{ik} P_k(t). \tag{7.1}$$

The coefficients A_{ik} are the transition probabilities. These probabilities are normalized as follows:

$$\begin{aligned} \sum_k P_k(t) &= 1, \\ \sum_i A_{ik} = \sum_k A_{kl} &= 1. \end{aligned} \tag{7.2}$$

We then have

$$P_i(t + \Delta t) - P_i(t) = \sum_k A_{ik} P_k(t) - \left(\sum_k A_{ki}\right) P_i(t). \tag{7.3}$$

We introduce the transition probabilities per unit time a_{ik} by

$$A_{ik} = (\Delta t) a_{ik}. \tag{7.4}$$

Then, in the limit $\Delta t \to 0$, Eq. (7.3) assumes the form

$$\frac{dP_i}{dt} = \sum_k (a_{ik} P_k - a_{ki} P_i). \tag{7.5}$$

This equation is called the master equation.

The master equation has a built-in irreversibility through Eq. (7.1). One can see the irreversibility clearly if one writes the inverse relation as

$$P_k(t) = \sum_i (A^{-1})_{ki} P_i(t + \Delta t). \tag{7.6}$$

The coefficients $(A^{-1})_{ki}$ do not have the meaning of transition probabilities. This is so because of a consistency requirement:

$$\sum_k A_{ik} (A^{-1})_{kj} = \delta_{ij}. \tag{7.7}$$

Since $1 \geqslant A_{ik} \geqslant 0$, all $(A^{-1})_{kj}$ cannot be positive or less than 1. Thus, if the coefficients $(A^{-1})_{kj}$ exist, they do not have the meaning of transition probabilities. Thus, we must recognize that the master equation has a built-in irreversibility.

Because of this important character of the master equation, we can prove that the H-function defined in terms of $P_i(t)$,

$$H = \sum_i P_i(t) \log P_i(t), \tag{7.8}$$

decreases with time. Indeed, using Eqs. (7.3) and (7.2), we find

$$\Delta H = \sum_i \Delta P_i \log P_i(t) \tag{7.9}$$

$$= \sum_i \left\{ \sum_k [A_{ik} P_k(t) - A_{ki} P_i] \right\} \log P_i(t)$$

$$= \sum_i \sum_k A_{ik} P_k(t) \log P_i(t) - \sum_i P_i \log P_i. \tag{7.10}$$

We now introduce a mathematical lemma which is analogous to (5.3).

LEMMA. *For two (probability) quantities satisfying*

$$1 \geqslant (P, Q) \geqslant 0, \tag{7.11}$$

the following holds:

$$F(P, Q) \equiv P(\log P) - P(\log Q) - P + Q \geqslant 0. \tag{7.12}$$

Proof. Since

$$F(P, P) = 0,$$

$$\left(\frac{\partial F}{\partial P}\right)_Q = \log\left(\frac{P}{Q}\right) \begin{cases} > 0 & \text{for } P > Q, \\ < 0 & \text{for } P < Q, \end{cases} \tag{7.13}$$

$F(P, Q)$ plotted as a function of P at a given Q is always positive, or zero when $P = Q$.

Using inequality (7.12) in Eq. (7.10) for $P_k = P$ and $P_i = Q$, we find

$$\Delta H \leq \sum_i \sum_k A_{ik}[P_k(\log P_k) - P_k + P_i] - \sum_i P_i \log P_i.$$

By applying the normalization conditions for P_i and A_{ik}, we find the right-hand side to be zero. Since ΔH corresponds to the time change in H for a time interval $\Delta t > 0$, we arrive at the following result in the limit $\Delta t \to 0$:

$$\frac{dH}{dt} \leqslant 0. \tag{7.14}$$

Thus, we conclude that if the probabilities P_i vary with time in accordance with the master equation, the H-function decreases with time until the equilibrium condition

$$a_{ik} P_k = a_{ki} P_i \tag{7.15}$$

is satisfied. This condition is the detailed balance condition for the master equation. The ratio

$$K_{ik} = \frac{a_{ik}}{a_{ki}} = \frac{P_i}{P_k} \tag{7.16}$$

determines the equilibrium constant.

Problem

Assuming in Eq. (7.16) that

$$K = e^{-\varepsilon/kT},$$

and also that n_i molecules are in the energy state ε, prove that the specific heat per molecule is given by

$$c = \frac{\varepsilon^2}{kT^2} \frac{e^{-\varepsilon/kT}}{(1 - e^{-\varepsilon/kT})^2}.$$

In concluding this section, we remark that a somewhat simplified version of the above proof of the nonpositiveness of ΔH can be given by using continuous variables. Let us introduce a transition probability $K(y, x)$, which connects two (positive) distribution functions $f(x)$ and $g(y)$ such that

$$g(y) = \int K(y, x) f(x)\, dx. \tag{7.17}$$

The transition probability may be normalized to one and assumed to be nonnegative:

$$\int K(y, x)\, dx = \int K(y, x)\, dy = 1, \tag{7.18}$$

$$K(y, x) \geqslant 0. \tag{7.19}$$

The normalization condition (7.18) ensures that for $g(y)$,

$$\int g(y)\, dy = 1. \tag{7.20}$$

Then one arrives at

$$\int g(y) \log g(y)\, dy \leqslant \int f(x) \log f(x)\, dx, \tag{7.21}$$

because

$$\int g(y) \log g(y)\, dy = \iint K(y, x) f(x)\, dx \log g(y)\, dy$$

$$\int f(x) \log f(x)\, dx = \int K(y, x)\, dy \int f(x) \log f(x)\, dx$$

and because that the difference of the right-hand sides of these expressions:

$$\iint K(y, x)[f(x) \log f(x) - f(x) \log g(y)]\, dx\, dy$$

cannot be negative. One can see this from Eq. (7.19) and also by modifying the integrand of this expression such that

$$\iint K(y, x) g(y) \left[\left(\frac{f(x)}{g(y)} \right) \log \left\{ \frac{f(x)}{g(y)} \right\} - \frac{f(x)}{g(y)} + 1 \right] dx\, dy.$$

Here $g(y) > 0$, and the factor in square brackets is nonnegative, due to Gibbs's inequality.

Note that the above proof is independent of time. However, by introducing the master equation or by using coarse graining processes temporal evolution of the entropy may be discussed (*8*).

Problems

1. Determine the volume of the phase space which is available to a harmonic oscillator with energy less than E.
2. The motion of a particle may be represented in the six-dimensional phase space (x, y, z, p_x, p_y, p_z). In this space, there are six independent variations

$\delta_1, \delta_2, \ldots, \delta_6$ which represent small displacements of the coordinates Note that

$$\frac{d(\delta \mathbf{r})}{dt} = \frac{\partial^2 \mathscr{H}}{\partial \mathbf{p}\, \partial \mathbf{r}} \cdot \delta \mathbf{r} + \frac{\partial^2 \mathscr{H}}{\partial \mathbf{p}\, \partial \mathbf{p}} \cdot \delta \mathbf{p}$$

$$\frac{d(\delta \mathbf{p})}{dt} = -\frac{\partial^2 \mathscr{H}}{\partial \mathbf{r}\, \partial \mathbf{r}} \cdot \delta \mathbf{r} - \frac{\partial^2 \mathscr{H}}{\partial \mathbf{r}\, \partial \mathbf{p}} \cdot \delta \mathbf{p}$$

and show that the volume element

$$V = \begin{vmatrix} \delta_1 x & \delta_1 y & \delta_1 z & \delta_1 p_x & \delta_1 p_y & \delta_1 p_z \\ \delta_2 x & \delta_2 y & \delta_2 z & \delta_2 p_x & \delta_2 p_y & \delta_2 p_z \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \delta_6 x & \delta_6 y & \delta_6 z & \delta_6 p_x & \delta_6 p_y & \delta_6 p_z \end{vmatrix},$$

corresponding to the displacements is not changed by the particle motion given by Eq. (1.1).

3. Starting with

$$\frac{d}{dt}(\mathbf{p} \cdot \mathbf{r}) = \frac{d\mathbf{p}}{dt} \cdot \mathbf{r} + \mathbf{p} \cdot \frac{d\mathbf{r}}{dt} = \mathbf{p} \cdot \frac{\partial \mathscr{H}}{\partial \mathbf{p}} - \mathbf{r} \cdot \frac{\partial \mathscr{H}}{\partial \mathbf{r}}$$

and assuming that both $\mathbf{p}$ and $\mathbf{r}$ are finite, show that

$$\left\langle \mathbf{p} \cdot \frac{\partial \mathscr{H}}{\partial \mathbf{p}} \right\rangle = \left\langle \mathbf{r} \cdot \frac{\partial \mathscr{H}}{\partial \mathbf{r}} \right\rangle$$

where $\langle \cdots \rangle$ represents the time average

$$\lim_{t \to \infty} \frac{1}{t} \int_0^t \cdots dt = \langle \cdots \rangle.$$

4. The distribution function of a dipole of moment μ in a uniform electric field of magnitude E may be written as

$$f(x) = e^{a+bx},$$

where $x = \cos\theta$ and $b = \mu E/kT$. Determine the constant a from the normalization condition for $f(x)$ that gives 2 in the domain $-1 \leqslant x \leqslant 1$ and evaluate the H-function as a function of b. Examine its sign.

5. If σ is a spin variable which takes on the values 1 or -1 and if h represents a magnetic field, the canonical distribution function for the spin may be expressed by

$$f(\sigma) = \frac{e^{\sigma h}}{\cosh h}.$$

Show that the H-function is

$$H(h) = 2(h \tanh h - \log \cosh h).$$

Examine the sign of H as a function of h.

6. The interaction potential between two dipoles at distance r may be given by

$$\phi = -\frac{\mu\mu'}{r^3}[2\cos\theta\cos\theta' - \sin\theta\sin\theta'\cos(\varphi - \varphi')].$$

The distribution function of these dipoles can approximately be given by

$$f(\omega, \omega') = A \exp\left\{\frac{\beta\mu\mu'}{r^3}[2\cos\theta\cos\theta' - \sin\theta\sin\theta'\cos(\varphi - \varphi')]\right\}$$

where ω and ω' are the solid angles specifying the orientations of the dipoles. Assuming that $\varepsilon_0\,(=\beta\mu\mu'/r^3)$ is small, determine A and evaluate H to order ε_0.

GENERAL REFERENCES

J. W. Gibbs, "Elementary Principles in Statistical Mechanics." Yale Univ. Press, New Haven, 1902.

A. I. Khinchin, "Mathematical Foundations of Statistical Mechanics." Dover, New York, 1949.

G. E. Uhlenbeck and G. W. Ford, "Lectures in Statistical Mechanics." Amer. Math. Soc., Providence, Rhode Island, 1963.

REFERENCES

1. J. von Neumann, *Proc. Nat. Acad. Sci. U.S.A.* **18**, 70, 263 (1932).
2. G. D. Birkoff, *Proc. Nat. Acad. Sci. U.S.A.* **17**, 656 (1931).
3. L. Hopf, "Ergoden Theorie." Springer, Berlin, 1937.
4. I. E. Farqhar, "Ergodic Theory in Statistical Mechanics." Wiley (Interscience), New York, 1964.
5. G. Sinai, *in* "Statistical Mechanics" (Thor Bak, ed.). Benjamin, New York, 1967.
6. A. Isihara and T. Wu, *Nuovo Cimento* **57B**, 83 (1968).
7. A. Isihara, *J. Phys. A (Proc. Phys. Soc.)* **2**, 539 (1968).
8. T. Ehrenfest and P. Ehrenfest, Sitzungsberichten der kaiserl. Akad der Wissenshaften in Wien. Mathem.—naturw. Klasse, **CXV**, abt. IIa, 89 (1906).

3

Partition Functions

Statistical mechanical methods of prime importance will be presented in this chapter. We shall study how the principles discussed in the previous chapter can be applied to systems in equilibrium. First, we shall derive the probability that a system in thermal equilibrium is in a particular energy state. Then, we shall find that the normalization of this probability gives a function of macroscopic variables which plays a fundamental role in our derivation of thermodynamic functions. The treatment of an ideal gas will serve as a guide throughout this chapter because of its simplicity and its connection with the definition of absolute temperature. Two other simple examples, on the elasticity of rubber and on lattice defects, are given at the end.

It is important to learn how partition functions can be constructed and used, and how a variational method can be employed effectively to determine probability distribution functions.

3.1. Boltzmann Statistics

We have learned in the previous chapter that macroscopic properties of a system can be deduced from the distribution function defined in the phase space. However, a quantum system cannot be represented by a phase point because the coordinates and the momenta of a particle cannot be determined completely at the same time (uncertainty principle). The method to be discussed in this section will be found useful even for a quantum system, although it was developed by Boltzmann and Gibbs before quantum mechanics.

Let us consider a system of identical particles with a discrete energy spectrum. That is, we take the energies to be discrete. For a continuous energy spectrum, we consider small energy cells obtained by partitioning. We ask: What is the probability that a particle in the system is in one of the energy levels?

To answer this question, let us assume that the number of particles N is very large. Let n_i be the number of particles in the ith state, with energy ε_i. Then,

$$N = \sum_i n_i, \tag{1.1}$$

where the summation is over all different states. If many states belong to the same energy, the summation extends over all these states. We shall find in later sections how the summation can be performed, especially for continuous energies.

The set of numbers n_i indicates the distribution of particles over the energy levels. The entire ensemble is described by this distribution. The ratio n_i/N represents the probability that a given particle has energy ε_i, because all the particles are identical.

In order to determine n_i for equilibrium, let us assume that there exists an average energy $\langle\varepsilon\rangle$ given by

$$\sum_i n_i \varepsilon_i = N\langle\varepsilon\rangle. \tag{1.2}$$

Furthermore, we assume the following:

(*i*) There are no restrictions on n_i other than Eqs. (1.1) and (1.2).
(*ii*) The permutations of the particles in different energy levels give rise to different configurations of the system, but those in the same energy level do not.

Under these assumptions, the number of different configurations of the system of N particles for a given set of numbers n_i is proportional to

$$W = \frac{N!}{\prod_i n_i!}. \tag{1.3}$$

W is the number of ways to distribute N particles among the energy levels. When Eq. (1.3) is adopted, we say that the original system obeys *Boltzmann statistics.*

The set of numbers n_i that maximizes W appears most frequently and corresponds to the most probable distribution. To determine such a distribution, we use the following mathematical formula.

STIRLING'S FORMULA (in the lowest order of approximation). *For a large integer n, the following approximation holds*:

$$\log n! = n(\log n) - n. \tag{1.4}$$

We give here a simple proof.[1]

$$\log n! = \sum_{x=1}^{n} \log x \sim \int_1^n \log x\, dx$$

$$= [x(\log x) - x]\Big|_1^n$$

$$\sim n(\log n) - n.$$

Using Stirling's formula in Eq. (1.3), we obtain the following expression:

$$\log W = N(\log N) - N - \sum_i [n_i(\log n_i) - n_i]. \tag{1.5}$$

We perform the variation of $\log W$ with respect to the n_i under the constraints

[1] Stirling's formula is derived as follows:

$$N! = \int_0^\infty e^{-t} t^N\, dt$$

$$= N^{N+1}(\exp -N)\int_{-1}^{\infty} (\exp -uN)(1+u)^N\, du \qquad [\text{by } t = N(1+u)]$$

$$= N^{N+1}(\exp\ \ N)\int_{-\infty}^{\infty} \exp\left(\frac{-Nv^2}{2}\right)\frac{du}{dv}\, dv \qquad \left[\text{by } (1+u)\exp -u = \exp\frac{-v^2}{2}\right]$$

$$\sim (2\pi)^{1/2} N^{N+1/2}(\exp -N)\left(\frac{du}{dv}\right)_0$$

$$= (2\pi N)^{1/2}\left(\frac{N}{e}\right)^N.$$

$$\log N! = N(\log N - 1) + \tfrac{1}{2}\log(2\pi N). \qquad (*)$$

Numerical values of $\log N!$ are given in the table below. As one sees, even for relatively small values of N, the two approximate formulas (1.4) and (*) are very good.

Approximations by Stirling's Formula

N	$\log N!$	$N(\log N) - N$	$(N+\frac{1}{2})(\log N) - N + \frac{1}{2}\log 2\pi$
5	4.8	3.0	4.8
10	15.1	13.0	15.1
20	42.3	39.9	42.3

of Eqs. (1.1) and (1.2), namely

$$\sum \delta n_i = 0, \tag{1.6}$$

$$\sum (\delta n_i)\varepsilon_i = 0. \tag{1.7}$$

Multiplying Eqs. (1.6) and (1.7) by the Lagrange multipliers A and $-\beta$, respectively, and adding the results to $\delta(\log W)$, we obtain

$$\sum_i (\delta n_i)(\log n_i - A + \beta\varepsilon_i) = 0. \tag{1.8}$$

For arbitrary variations δn_i, the terms in the bracket must vanish. Therefore, we obtain

$$n_i = N\frac{e^{-\beta\varepsilon_i}}{Z}, \tag{1.9}$$

where Z is a constant related to A. It can be determined from the normalization condition (1.1) as follows:

$$Z = \sum_i e^{-\beta\varepsilon_i}. \tag{1.10}$$

Z is the sum of the *Boltzmann factors* $e^{-\beta\varepsilon_i}$ over all states.

If ε_i is the kinetic energy of a gas particle, the expression for n_i becomes the Maxwell distribution function because Eq. (1.2) determines β as $1/kT$ when it is applied to an ideal gas. Note that Eq. (1.9) can be used for the cases where ε_i includes potential energies due to external fields.

It is important to recognize that Eq. (1.5) is essentially

$$\log W = -\sum_i n_i \log n_i. \tag{1.11}$$

Comparing this with Eqs. (3.1) and (3.10) of Chapter 2, we may conclude that, apart from a constant, this expression gives the entropy if it is multiplied by the Boltzmann constant k:

$$S = k \log W. \tag{1.12}$$

Therefore, we have determined n_i based on a variational method which maximized the entropy. Equation (1.12) is called the *Boltzmann relation.*

3.2. Partition Function

The normalization constant Z defined by Eq. (1.10) is called the *partition function*, also called the *sum over states*. The former name corresponds to the partition of the ensemble into sets of systems in different energy levels, and the latter name refers simply to the summation over i. Indeed, the

partition function is obtained by summing Boltzmann factors over all states. If the system has continuous energies, the summation is replaced by integration over the phase space. Note that Z is obtained as a result of summing over all microscopic states and is not a function of an individual state. As we shall discuss shortly, it is a function of macroscopic variables such as volume and temperature. Through Z, we connect macroscopic quantities with microscopic states.

The method of the previous section is applicable to a real system such as a gas in which particles are interacting. Let us introduce an *ensemble* of identical gas systems composed of the same number of particles. Figure 3.1

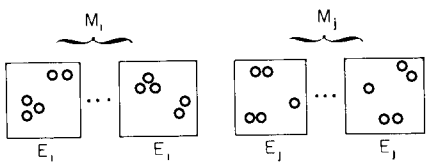


Fig. 3.1. An ensemble of gas systems.

illustrates each system in such an ensemble by a box. There are M_i boxes (systems) in the energy state E_i, there are M_j boxes of energy E_j, etc. Since there is no interaction between these boxes, we can use the method of the previous section to prove that the probability of finding a given gas system in the energy state E_i is

$$P_i = \frac{e^{-\beta E_i}}{Z}. \tag{2.1}$$

If M is the total number of boxes, we have

$$M_i = P_i M. \tag{2.2}$$

The normalization constant Z is the partition function of the system:

$$Z = \sum_i e^{-\beta E_i}. \tag{2.3}$$

The distribution described by Eq. (2.1) is identical to the canonical distribution of Eq. (3.4) of Chapter 2. The partition function depends on $\beta = 1/kT$ and the total number of particles in the system. Since the energy E_i can depend on external forces, Z can also depend on these forces. In addition, Z can depend on the volume V of the system because the states i can be determined by the entire volume.

The name *canonical ensemble* is used for an ensemble of identical systems with the above probability distribution in equilibrium with a heat bath. Gibbs obtained the distribution function P_i given by Eq. (2.1) by considering

such an ensemble. Therefore, this method of deriving the distribution function for a canonical ensemble is called *Gibbs's ensemble method.*

Using the expression for P_i, we can evaluate *canonical ensemble averages.* For instance, the average energy is

$$U = \sum_i \frac{E_i e^{-\beta E_i}}{Z}. \tag{2.4}$$

Caution is necessary in identifying this average energy. If E_i includes the energy due to the motion of a gas system as a whole, it is not the internal energy. Assuming that there is no center-of-mass motion, we take Eq. (2.4) as the statistical definition of the internal energy.

The establishment of the relations between statistical averages and thermodynamic functions can be made by the following procedure. We note that Eq. (2.4) can be obtained by differentiating $\log Z$ with respect to $-\beta$. Remembering that Z is a function of the temperature and total volume V, we take the total derivative of $kT \log Z$:

$$-d(kT \log Z) = -kT\left(\frac{\partial \log Z}{\partial T} + \frac{1}{T}\log Z\right) dT - \frac{\partial (kT \log Z)}{\partial V} dV. \tag{2.5}$$

On the other hand, the second law of thermodynamics gives

$$dU = T\,dS - p\,dV, \tag{2.6}$$

where U is the internal energy, S is the entropy, and p is the pressure. Thus, the Helmholtz free energy defined by

$$F = U - TS \tag{2.7}$$

gives

$$dF = -S\,dT - p\,dV \tag{2.8}$$

$$= -\left(\frac{U-F}{T}\right) dT - p\,dV. \tag{2.9}$$

We compare Eq. (2.5) with Eq. (2.9). Observe that the statistical average energy given by Eq. (2.4) is

$$U = -\left(\frac{\partial \log Z}{\partial \beta}\right)_V = \left(\frac{\partial \log Z}{\partial T}\right)_V kT^2. \tag{2.10}$$

If we consider U as the thermodynamic internal energy, Eqs. (2.5) and (2.9)

become identical, with the following correspondences:

$$F = -kT \log Z, \tag{2.11}$$

$$p = -\left(\frac{\partial F}{\partial V}\right)_T, \tag{2.12}$$

$$= kT\left(\frac{\partial(\log Z)}{\partial V}\right)_T. \tag{2.13}$$

These relations between statistical averages and thermodynamic quantities are in agreement with what we have derived in Chapter 2. The establishment of these relations may not be very convincing, but it is partly based on statistical definitions and partly on the analytic behavior of thermodynamic functions. As we proceed further, these relations will become more clearly understood.

Problems

1. Find that β must be inversely proportional to T by applying Eq. (2.4) to an ideal gas.
2. Derive the following formulas:

$$U = -T^2\left[\frac{\partial(F/T)}{\partial T}\right]_V, \tag{2.14}$$

$$C_V = -\frac{\partial}{\partial T}\left[T^2\frac{\partial(F/T)}{\partial T}\right]_V. \tag{2.15}$$

3. Using Eqs. (2.3) and (2.13), find the equation of state of an ideal gas.
4. Derive the equation of state and the specific heat of an ideal gas mixture.

3.3. Gibbs's Paradox

In the previous section, the partition function was introduced assuming that the energy levels are discrete. The partition functions for classical systems with continuous energy levels may be obtained by taking a suitable limit. For instance, if the Hamiltonian $\mathscr{H}(\mathbf{p}, \mathbf{r})$ of a classical system is known as a function of $\mathbf{p}$ and $\mathbf{r}$, the summation of $e^{-\beta E_i}$ over states may be replaced by the integration of $e^{-\beta\mathscr{H}}$ over the phase space $(\mathbf{p}, \mathbf{r})$. In many cases, the

Hamiltonian is composed of the kinetic energy and the potential energy which is a sum of potentials between two molecules:

$$\mathscr{H} = \sum_i \frac{p_i^2}{2m} + \sum_{i<j}\phi(r_{ij}). \tag{3.1}$$

Here, r_{ij} is the relative distance between the molecules. The partition function of Eq. (2.3) may be evaluated from

$$Z = A\int e^{-\beta\mathscr{H}}\, d\mathbf{p}^N\, d\mathbf{r}^N, \tag{3.2}$$

where A is a proportionality constant to keep the dimension of the partition function unity as in Eq. (2.3). In what follows, we discuss how A can be determined.

The partition function plays an important role in the evaluation of thermodynamic functions. Since the Helmholtz free energy is porportional to the logarithm of the partition function as shown in Eq. (2.11), the proportionality constant A need not be determined as far as the derivatives of the free energy are concerned. It must be determined, however, when the absolute magnitude of the free energy or the entropy comes into consideration. To find A, let us treat an ideal gas. The partition function is given by

$$Z_0 = A\left[\int \exp\left(-\frac{\beta p^2}{2m}\right) d\mathbf{r}\, d\mathbf{p}\right]^N = A(2\pi mkT)^{3N/2}V. \tag{3.3}$$

Therefore, we find the free energy, the internal energy and the entropy as

$$\begin{aligned} F_0 &= -NkT\log[(2\pi mkT)^{3/2}V] - kT\log A, \\ U_0 &= \tfrac{3}{2}NkT, \\ S_0 &= Nk\{\tfrac{3}{2} + \log[(2\pi mkT)^{3/2}V]\} + k\log A, \end{aligned} \tag{3.4}$$

where the subscript zero indicates that the quantities are for an ideal gas.

Now let us consider two identical ideal gas systems of equal volume V with the same number of molecules N at the same temperature T. If we combine these two systems by removing a diaphragm between them, we should have twice the free energy (or the entropy) given by Eqs. (3.4). On the other hand, if from the beginning we have $2N$ molecules in the volume $2V$, the corresponding free energy (or entropy) can be obtained by replacing N in Eqs. (3.4) by $2N$ and V by $2V$. Thus, these two processes yield the two equations

$$F_0 = -2NkT\log[(2\pi mkT)^{3/2}V] - 2kT\log A \tag{3.5'}$$

$$F_0 = -2NkT\log[(2\pi mkT)^{3/2}2V] - kT\log A \tag{3.5''}$$

Equations (3.5′) and (3.5″) are not compatible. This contradiction is called *Gibbs's paradox*. Gibbs solved this paradox by an ingenious insight. He argued that permutations of identical molecules should not yield new configurations, and the partition function should have the factor

$$A \sim \frac{1}{N!}. \tag{3.6}$$

The reader may check that this factor makes Eqs. (3.5′) and (3.5″) identical: Gibbs's paradox is removed by the factor $1/N!$. It is important to notice that this consideration meets the quantum mechanical concept that identical particles are indistinguishable.

In this respect, let us discuss another factor to be included in A which is due to a quantum mechanical requirement. Our discussions have so far been based on classical mechanics, and the states of our system are represented by phase points. According to quantum mechanics, however, Heisenberg's uncertainty relation requires

$$d\mathbf{p}\, d\mathbf{r} > h^3. \tag{3.7}$$

Therefore, h^3 is the smallest possible phase-space volume per particle. In view of this, we introduce h^{3N} for the unit of phase-space volume for N-particle systems and express the classical partition function as follows:

$$Z_N = \frac{1}{h^{3N}N!}\int e^{-\beta\mathscr{H}}\, d\mathbf{p}^N\, d\mathbf{r}^N. \tag{3.8}$$

This expression is dimensionless, as desired.

The expressions for the free energy, the internal energy, and the entropy of an ideal gas follow from Eq. (3.8):

$$\begin{aligned} F_0 &= -NkT\log\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2} v\right], \\ U_0 &= \tfrac{3}{2}NkT, \\ S_0 &= Nk\left\{\frac{3}{2} + \log\left[\left(\frac{2\pi mkT}{h^2}\right)^{3/2} v\right]\right\}. \end{aligned} \tag{3.9}$$

These quantities are *extensive variables*: They are proportional to the total number N.

Because of the uncertainty principle, we take the approach of not distinguishing phase points inside the elementary volume h^{3N} but of considering them as having the same probability. This is an example of the *a priori probability assumption* in statistical mechanics. We may attach a small volume h^{3N} to a given phase point and consider the motion of the volume. The trajectory will eventually cover the whole energy surface. Then, a time average over the trajectory may be replaced by a phase-space average (ergodicity).

3.4. Grand Ensemble

The canonical ensemble discussed in Section 3.2 can be generalized to a more general ensemble in which the number of particles varies. Let us consider a collection of M similar systems, $a_N{}^k$ of which have N particles in the energy state $E_N{}^k$. The set of numbers $a_N{}^k$ represents a distribution of these systems in various energies and numbers. For a given set of the $a_N{}^k$, the ensemble takes on different configurations, the total number of which is given by

$$W = \frac{M!}{\prod_{N,k} a_N{}^k!}. \tag{4.1}$$

The quantity W is the weight of the ensemble for a given distribution characterized by the numbers $a_N{}^k$. The larger this weight is, the more frequently will the distribution appear.

Let us seek the distribution that maximizes W. To specify the distribution, we introduce the following constraints: The total number of systems, the total number of particles, and the total energy of the ensemble are constants:

$$\begin{aligned} \sum_{N,k} a_N{}^k &= M, \\ \sum_{N,k} N a_N{}^k &= N_0, \\ \sum_{N,k} E_N{}^k a_N{}^k &= E_0. \end{aligned} \tag{4.2}$$

As in Section 3.1, such a distribution may be considered to be the most probable one and as corresponding to equilibrium.

Taking the variation of $\log W$ with respect to the $a_N{}^k$ with the Lagrange multipliers ν, ζ, and β, we find

$$\sum_{N,k} [(\log a_N{}^k) + \nu + N\zeta + \beta E_N{}^k]\, \delta a_N{}^k = 0. \tag{4.3}$$

Since the $\delta a_N{}^k$ are arbitrary, the terms in the bracket must vanish, giving

$$\begin{aligned} a_N{}^k &= \exp(-\nu - \zeta N)\exp(-\beta E_N{}^k). \\ &= \frac{M}{\Xi}\lambda^N \exp(-\beta E_N{}^k). \end{aligned} \tag{4.4}$$

Equation (4.4) gives the most probable distribution; β is common to all the systems and, as in Chapter 2, can be identified as $1/kT$; $\lambda = e^{-\zeta}$; and

$$\begin{aligned} \Xi &= \sum_{N,k} \lambda^N \exp(-\beta E_N{}^k) \\ &= \sum_N \lambda^N Z_N. \end{aligned} \tag{4.5}$$

The probability that the system has N particles and is the energy level $E_N{}^k$ is represented by $a_N{}^k/M$. Thus, from Eq. (4.4), we introduce the probability

$$P_N{}^k = \frac{\lambda^N \exp(-\beta E_N{}^k)}{\Xi}. \tag{4.6}$$

For a fixed N, $P_N{}^k$ is of the form of the canonical distribution. λ^N is the weight for having N particles. The collection of systems with the probability given by Eq. (4.6) is called the *grand ensemble* and Ξ is the *grand partition function.* The grand partition function is very useful in treating many-body systems when the number of particles fluctuates.

If we sum $P_N{}^k$ over all energy states k, we have

$$P_N = \frac{\lambda^N Z_N}{\Xi}. \tag{4.7}$$

Further summation over N gives

$$\sum_N P_N = 1. \tag{4.8}$$

Each term in this summation is positive and P_N is the probability of the system of N particles in the grand ensemble.

By taking the total derivative of Eq. (4.5), we find

$$\begin{aligned} d(\log \Xi) &= \sum_N N P_N \, d(\log \lambda) + \sum_{N,k} (-E_N{}^k) P_N{}^k \, d\beta - \beta \sum_{N,k} P_N{}^k \, dE_N{}^k \\ &= \langle N \rangle \, d(\log \lambda) - U \, d\beta + \beta p \, dV, \end{aligned} \tag{4.9}$$

where

$$\begin{aligned} \langle N \rangle &= \sum_N N P_N, \\ U &= \sum_{N,k} E_N{}^k P_N{}^k, \\ p &= -\sum_{N,k} \frac{\partial E_N{}^k}{\partial V} P_N{}^k \end{aligned} \tag{4.10}$$

are the grand ensemble averages to be identified respectively as the average number of particles in the system, the internal energy, and the pressure. In the last equation, it is assumed that changes in the energies at constant temperature and λ are due to changes in volume. In the case when the system is subjected to an external force, a similar relation will be obtained between the force and the corresponding strain.

With use of Eq. (4.9), Eqs. (4.10) may be rewritten in terms of the grand partition function Ξ:

$$\begin{aligned} \langle N \rangle &= \left[\frac{\partial(\log \Xi)}{\partial(\log \lambda)}\right]_{\beta, V}, \\ U &= -\left[\frac{\partial(\log \Xi)}{\partial \beta}\right]_{\lambda, V}, \\ p &= \frac{1}{\beta}\left[\frac{\partial(\log \Xi)}{\partial V}\right]_{\lambda, \beta}. \end{aligned} \tag{4.11}$$

The last equation is based on the following:

$$\begin{aligned} \frac{1}{\beta}\left[\frac{\partial(\log \Xi)}{\partial V}\right]_{\lambda, \beta} &= \frac{1}{\beta} \frac{\sum_N \left(\lambda^N \frac{\partial Z_N}{\partial V}\right)}{\sum_N \lambda^N Z_N} \\ &= \frac{\sum_N p_N \lambda^N Z_N}{\sum_N \lambda^N Z_N}, \end{aligned}$$

where p_N is the pressure for a system with N particles. We shall find in the next section that the pressure p is more directly given by the grand partition function.

3.5. Relation between the Canonical and Grand Canonical Partition Functions

The grand ensemble method is applicable to a system of a definite number of particles. In this case, the grand-ensemble average number of particles must be the given total number of particles. In the grand ensemble method, we consider fluctuations in the number of particles, but the fluctuations can be very small. This consideration leads us to expect that in the series defining the grand partition function,

$$\Xi(T, V; \lambda) = \sum_N \lambda^N Z_N, \tag{5.1}$$

there is a maximum term corresponding to the most probable number and that its contribution is overwhelmingly large compared with those from other terms. In Eq. (5.1), $Z_N(T, V)$ is the partition function for a system of N particles, λ is a parameter to be determined later, and the summation goes over all possible N values.

Based on the above consideration, let us find the relation between the canonical and the grand canonical partition functions. According to

Cauchy's theorem, the partition function for a system of N particles can be obtained from the grand partition function by evaluating the contour integral

$$Z_N = \frac{1}{2\pi i}\oint \frac{\Xi}{\lambda^{N+1}}\, d\lambda, \tag{5.2}$$

where λ is now a complex variable.

For large N, the integral may be evaluated by the method of steepest descent. The application of this method is analogous to that of a variational method in the previous section.

From Eq. (2.11), we have

$$Z_N = e^{-\beta F}, \tag{5.3}$$

where F is the Helmholtz free energy. Because the Helmholtz free energy F is proportional to N for a system of N particles, it is convenient to introduce

$$\Xi = e^{-\beta\Omega}, \qquad \lambda = e^{\beta\mu}. \tag{5.4}$$

Ω is called the *grand potential.*

Equation (5.2) becomes

$$Z_N = \frac{1}{2\pi i}\oint \frac{e^{-\beta(\Omega+N\mu)}}{\lambda}\, d\lambda. \tag{5.5}$$

The saddle point is determined from the condition that the first derivative of the exponential function in the integrand of Eq. (5.5) vanishes:

$$\left[\frac{\partial\Omega}{\partial(\log\lambda)}\right]_{T,V} + N\left[\frac{\partial\mu}{\partial(\log\lambda)}\right]_{T,V} = 0,$$

or

$$-\beta\left[\frac{\partial\Omega}{\partial(\log\lambda)}\right]_{T,V} = N. \tag{5.6}$$

This is simply the first of Eqs. (4.11).

Using the value of λ (or μ) determined at the saddle point, the first approximation for Z_N is given by

$$Z_N = e^{-\beta(\Omega+N\mu)}. \tag{5.5'}$$

Comparing this with Eq. (5.3), we obtain

$$-\Omega = N\mu - F. \tag{5.7}$$

Therefore,

$$-d\Omega = N\,d\mu + p\,dV + S\,dT. \tag{5.8}$$

On the other hand, the Gibbs free energy used in thermodynamics is defined by

$$G(T, P) = F + pV \tag{5.9}$$

$$= U - TS + pV$$

$$= N\mu, \tag{5.10}$$

where μ is called the chemical potential and is the Gibbs free energy per molecule. Using Eq. (5.9), we find

$$dG = -S\,dT + V\,dp \tag{5.11}$$

$$dG = N\,d\mu. \tag{5.12}$$

From Eqs. (5.9) and (5.11), we obtain the relation

$$d(pV) = N\,d\mu + p\,dV + S\,dT. \tag{5.13}$$

The right-hand side of this equation is identical to that of Eq. (5.8). Thus, we establish the relation

$$-\beta\Omega = \log\Xi = \frac{pV}{kT}. \tag{5.14}$$

At the same time, μ defined by Eq. (5.4) is identified as the chemical potential. The λ appearing in the same equation is called the *absolute activity*, which is often confused with the *fugacity*.[2] Equation (5.6) can be rewritten as

$$N = \frac{\sum_M M\lambda^M Z_M}{\sum_M \lambda^M Z_M}. \tag{5.15}$$

From this expression, $\lambda^M Z_M$ can be interpreted as the relative probability at temperature T and volume V of finding the system with M particles. The total ensemble of systems having these probabilities is the grand ensemble. Equation (5.15) gives the grand-ensemble average number of particles.

---

[2] The words *fugacity* and *activity* are confused in literature. Correctly, fugacity should be used for a quantity having the dimension of a gaseous pressure. Thus if

$$z = a\exp(\mu/kT),$$

where μ is the chemical potential; and if a is chosen at each temperature such that z tends to the partial pressure at infinite dilution then it is fugacity. If this is not the case, z should be called activity. In particular, if a is put equal to 1 independent of the temperature z is called *absolute activity*.

Let us treat a classical perfect gas by the grand partition function method. The absolute activity for this case is denoted by λ_0. We have the grand partition function

$$\begin{aligned}\Xi_0 &= \sum_N \frac{\lambda_0^N}{N!}\left[\frac{1}{h^3}\int \exp\left(\frac{-p^2}{2mkT}\right) d\mathbf{p}\, d\mathbf{r}\right]^N \\ &= \sum_N \frac{1}{N!}\left[\lambda_0 V\left(\frac{2\pi mkT}{h^2}\right)^{3/2}\right]^N \\ &= \exp V z_0 . \end{aligned} \tag{5.16}$$

The quantity z_0 appearing in the last expression is called the *activity*. It is related to λ_0 by

$$z_0 = \lambda_0\left(\frac{2\pi mkT}{h^2}\right)^{3/2}. \tag{5.17}$$

Thus,

$$\frac{p}{kT} = z_0, \tag{5.18}$$

and

$$\begin{aligned} z_0 &= n; \\ \lambda_0 &= \frac{n}{(2\pi mkT/h^2)^{3/2}}. \end{aligned} \tag{5.19}$$

The quantity $h/(2\pi mkT)^{1/2}$ is called the *thermal de Broglie wavelength* and will be discussed later in Section 4.3. Note in Eqs. (5.19) that z_0 is equal to n, and λ_0^{-1} is the partition function per particle for a volume $v = 1/n$.

Returning to the general case where interactions may be present, we obtain from Eqs. (5.4), (5.7), and (5.14) that

$$\lambda = \frac{e^{\beta pv}}{e^{-\beta f}}, \tag{5.20}$$

where f and v are the Helmholtz free energy and the volume per particle,

$$f = \frac{F}{N}, \qquad v = \frac{V}{N}. \tag{5.21}$$

$e^{-\beta f}$ is the partition function per particle. In particular, in the case of an ideal gas, the partition function for N molecules is

$$Z_N{}^0 = \frac{1}{N!}\frac{V^N}{h^{3N}}(2\pi mkT)^{3N/2}. \tag{5.22}$$

Thus, using Stirling's formula, we obtain

$$\begin{aligned}(Z_N{}^0)^{1/N} &= ev\left(\frac{2\pi mkT}{h^2}\right)^{3/2} \\ &= e^{-\beta f_0}. \end{aligned} \tag{5.23}$$

On the other hand, for an ideal gas, Eq. (5.20) gives

$$\lambda_0 = \frac{e}{e^{-\beta f_0}} = \frac{e}{(Z_N{}^0)^{1/N}} = \frac{n}{(2\pi mkT/h^2)^{3/2}}. \tag{5.24}$$

This result confirms Eq. (5.19). Taking the ratio of λ and λ_0 given by Eqs. (5.24) and (5.20), we find

$$\beta(\mu - \mu_0) = \beta(f + pv) - (\beta f_0 + 1). \tag{5.25}$$

μ is the chemical potential defined in Eq. (5.10).

In the above calculation, $N!$ plays an important role. This factor has been introduced from the quantum mechanical consideration of indistinguishability of particles. In the case when the magnitudes of the thermodynamic functions come into consideration, the factor $N!$ is always important.

Problem

Derive the equation of state for an ideal gas mixture using the grand partition function.

3.6. Fluctuations

We have seen that the first derivatives of the partition function or the grand partition function give average quantities. The second derivatives also yield useful results. Let us differentiate with respect to λ at constant T and V:

$$e^{\beta\Omega}\sum \lambda^N Z_N = 1, \tag{6.1}$$

where

$$-\beta\Omega = \log \Xi;$$

then,

$$e^{\beta\Omega}\sum_N \lambda^N Z_N\left[N - \frac{\partial(\log \Xi)}{\partial(\log \lambda)}\right]\delta(\log \lambda) = 0. \tag{6.2}$$

Remembering the probability character of each term of the left-hand side of Eq. (6.1), we get

$$\left[\frac{\partial(\log \Xi)}{\partial(\log \lambda)}\right]_{T,V} = \langle N \rangle, \tag{6.3}$$

where $\langle N \rangle$ means the grand ensemble average of N.

Differentiating Eq. (6.2) again and using Eq. (6.3), we find

$$e^{\beta\Omega} \sum_N \lambda^N Z_N \left[N - \frac{\partial(\log \Xi)}{\partial(\log \lambda)}\right]^2 (\delta \log \lambda)^2 + e^{\beta\Omega} \sum_N \lambda^N Z_N \left[-\frac{\partial^2(\log \Xi)}{\partial(\log \lambda)^2}\right] (\delta \log \lambda)^2 = 0. \tag{6.4}$$

Thus,

$$\left[\frac{\partial^2(\log \Xi)}{\partial(\log \lambda)^2}\right]_{T,V} = \langle (N - \langle N \rangle)^2 \rangle \tag{6.5}$$

i.e., the second derivative $[\partial^2(\log \Xi)/\partial(\log \lambda)^2]_{T,V}$ gives the average fluctuation of N about its mean value.

Similarly, we find

$$U = \langle E \rangle = -\left[\frac{\partial(\log \Xi)}{\partial \beta}\right]_{\lambda,V}, \qquad \langle (E - \langle E \rangle)^2 \rangle = \left[\frac{\partial^2(\log \Xi)}{\partial \beta^2}\right]_{\lambda,V}. \tag{6.6}$$

The average fluctuations are generally small. For example, in the case of an ideal gas, we have $-\beta\Omega_0 = \lambda V(2\pi mkT)^{3/2}/h^3$. Therefore,

$$\frac{\langle (N - \langle N \rangle)^2 \rangle}{\langle N \rangle^2} = \frac{1}{\langle N \rangle}, \qquad \frac{\langle (E - \langle E \rangle)^2 \rangle}{\langle E \rangle^2} = \frac{5}{3\langle N \rangle}. \tag{6.7}$$

These equations show that the average fluctuations are of the order of $1/\langle N \rangle$. Exactly the same results can be obtained using the canonical partition function.

Problems

1. Find the average fluctuation of the pressure of an ideal gas.
2. Confirm Eq. (4.9) by using Eq. (5.13).
3. Show that the specific heat at constant volume is positive.

3.7. The Elasticity of Rubber

Slightly vulcanized rubber shows a remarkable extensibility, which may reach up to 300% within the elastic range. This contrasts with about 0.1% for other substances. In addition, the elasticity of rubber is characterized by the following:

1. Hooke's law does not hold for rubber. The stress–strain curve is not the straight line characteristic of Hookean elasticity but is S-shaped (Fig. 3.2).

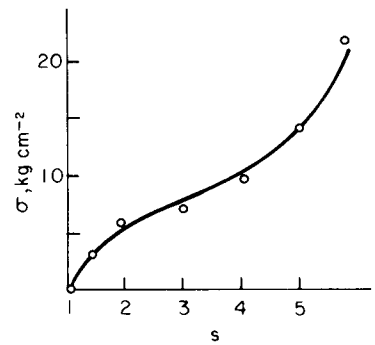


Fig. 3.2. Stress–strain curve of slightly vulcanized rubber.

2. The elasticity of rubber is enhanced at higher temperatures, while usual substances lose elasticity: the elastic constant at a fixed elongation is proportional to the absolute temperature.
3. An adiabatic extension raises the temperature of rubber.
4. The volume change upon elongation is very small; the Poisson ratio is approximately 0.5.

Vulcanized rubber is a solid and maintains its shape, while nonvulcanized rubber flows. Property 2 resembles that of gas pressure, which also increases with temperature. Property 4 suggests a similarity of rubber with liquids.

A rubber molecule has a large molecular weight, of the order of 500,000 ~ 1,000,000. Natural rubber, i.e., polyisoprene $(C_5H_8)_n$, consists of a large number n of isoprene (C_5H_8) units. The degree of polymerization n is of the order 10,000. Because of the largeness of n, one applies statistical mechanics to the rubber molecule.

Raw rubber does not have the remarkable elasticity noted above. Vulcanization is considered to bind rubber molecules together to form a three-dimensional network so that no plastic flow takes place. In view of the free internal rotation of rubber molecules, let us introduce a simple model in which rubber chains between two successive vulcanized points (sulfur bridges)

are composed of units effectively independent of each other, called segments. These segments have a constant length b. They can rotate freely without structural hindrance. A segment may consist of several isoprene units.

Let us consider a representative chain between two successive vulcanization points at $\mathbf{r}_a$ and $\mathbf{r}_b$. We choose a coordinate system attached to this chain such that the origin is at $\mathbf{r}_b$ and the polar axis is in the direction of $\mathbf{r}_a - \mathbf{r}_b$. Let $f(\theta, \varphi) \sin\theta \, d\theta \, d\varphi$ be the number of segments having polar angles between θ and $\theta + d\theta$ and φ and $\varphi + d\varphi$. According to Eqs. (1.11) and (1.12), the free energy may be written as follows (*1*):

$$F = U + kT \sum \int f(\theta, \varphi)[\log f(\theta, \varphi)] \sin\theta \, d\theta \, d\varphi. \tag{7.1}$$

The first term U is the internal energy. The summation goes over all the chains in the sample.

In view of Property 4, let us assume that U is a function of volume only. We determine the distribution function $f(\theta, \varphi)$ by minimizing the second term, which is due to the entropy, under the constraints

$$\begin{aligned} \int f(\theta, \varphi) \sin\theta \, d\theta \, d\varphi &= N, \\ \int b \cos\theta \, f(\theta, \varphi) \sin\theta \, d\theta \, d\varphi &= r. \end{aligned} \tag{7.2}$$

The first constraint means that the total number of segments between the vulcanization points is N. The second constraint means that the projection of the segmental vectors onto the polar axis is equal to the end-to-end distance $r = |\mathbf{r}_b - \mathbf{r}_a|$ of the chain.

The values of N and r vary from chain to chain. The directions of the chains may also vary. The summation in Eq. (7.1) goes over all these variations. Using the method of Lagrange multipliers, we arrive at

$$\lambda f(\theta, \varphi) = e^{\mu \cos\theta}, \tag{7.3}$$

where λ and μ are Lagrange multipliers. Equations (7.2) give

$$\begin{aligned} \lambda N &= 4\pi \frac{\sinh\mu}{\mu}, \\ \frac{\lambda r}{b} &= 4\pi \left(\frac{\cosh\mu}{\mu} - \frac{\sinh\mu}{\mu^2} \right). \end{aligned} \tag{7.4}$$

From these two equations, one can obtain λ and μ as functions of r/Nb, which is the ratio of the end-to-end distance to the maximum length of the chain. In terms of λ and μ, the free energy is given by

$$F = U + kT \sum \left(\frac{\mu r}{b} - N \log \lambda \right). \tag{7.5}$$

Let us now evaluate the stress. Imagine a unit cube of a rubber sample. Upon stretching in the z direction, the cube will be a rectangular solid of side lengths s_x, s_y, and s_z. In view of Property 4, we require

$$s_z s_x^2 = s_z s_y^2 = 1. \tag{7.6}$$

The stress σ_z is obtained by differentiating the free energy with respect to s_z. From $U = U(s_x s_y s_z)$, we have $\partial U/\partial s_z = s_x s_y U'$, where U' is the derivative of U with respect to volume. If the stress is in the z direction only, $\sigma_y = \sigma_x = 0$. Using these relations, one can eliminate U' from the equation for $\partial F/\partial s_z$. For small elongation, we obtain

$$\frac{\sigma_z}{kT} = \sum N \left[\left(\frac{r_0}{Nb} \right)^2 \left(s_z - \frac{1}{s_z^2} \right) + \frac{9}{25} \left(\frac{r_0}{Nb} \right)^4 \left(s_z^3 + \frac{1}{3} - \frac{4}{3 s_z^3} \right) + \cdots \right], \tag{7.7}$$

where r_0 is the end-to-end distance in the absence of the stress and σ_z is the stress per unit of original cross section.

The stress evaluated to higher orders reproduces the experimental S-curve (Fig. 3.2). The elastic constant at zero elongation assumes the form

$$E_0 = 3\rho \left(\frac{RT}{m} \right) \left\langle \left(\frac{r_0}{Nb} \right)^2 \right\rangle, \tag{7.8}$$

where ρ is the density, m is the molecular weight of the segments, and the angular brackets mean a statistical average; ρ is of the order of unity, $\langle (r_0/Nb)^2 \rangle$ is approximately $(1/8)^2$. The value $m = 120$ gives E_0 of order 10 kg cm^{-2} at $T = 300°$K, which is correct in order of magnitude. The proportionality of E_0 with the absolute temperature T also agrees with experiment.

The above treatment can be refined by taking into consideration the displacements of vulcanization points. According to the present model, the configuration of a rubber chain is determined by the ratio r_0/Nb. The chain can take on more configurations at shorter end-to-end distances. It loses entropy with extension. This explains the temperature increase of a rubber sample with adiabatic extension.

The above statistical evaluation of the configurations of a rubber molecule is improved if correlations of segments are taken into consideration. Even if these repeating units are replaced simply by rigid spheres, the configurational statistics becomes very complicated. We confront what is known as the *self-avoiding walk problem*. This problem is important because a similar problem arises in Ising lattice statistics. We shall discuss these problems in later chapters.

3.8. Lattice Defects

Crystals usually include lattice defects such as vacancies, interstitial atoms, or foreign atoms. Their effects are particularly noticeable in such crystals as alkali halides, silver halides, germanium, silicon, copper, and tin. The imperfections play important roles in the conductivity of semiconductors and of some good conductors, in the color of crystals, in luminescense, and in the mechanical properties of crystals (*2*).

3.8.1. Schottky Defects

The simplest type of defect occurs when crystal atoms move out to the surface, leaving vacancies called *Schottky defects*. If ε_s is the energy required to move a lattice atom to the surface, and if there are n vacancies distributed randomly in N lattice sites, the partition function of the lattice associated with the defects is

$$Z = \frac{N!}{(N-n)!\,n!}\, e^{-n\varepsilon_s/kT}, \tag{8.1}$$

where the first factor represents the number of configurations that the system can take on for given n and ε_s. Since there is only one energy value ε_s, there is no summation over energy states. The free energy is given by

$$F = n\varepsilon_s - kT \log \frac{N!}{(N-n)!\,n!}; \tag{8.2}$$

n is determined by the condition of minimum free energy:

$$\frac{\partial F}{\partial n} = 0. \tag{8.3}$$

For $N \gg n$, we find

$$n = Ne^{-\varepsilon_s/kT}. \tag{8.4}$$

If lattice vacancies are formed in pairs, as often occurs, the distribution of vacancies is given by

$$n = Ne^{-\varepsilon_p/2kT}; \tag{8.5}$$

where ε_p is the pair-formation energy. In NaCl, the order of magnitude of ε_p is 2.02 eV at room temperature, giving $n \sim 10^{-6}$ cm^{-3}.

3.8.2. Frenkel Defects

Sometimes, regular lattice particles move to interstitial positions and form *Frenkel defects*. If ε is the energy required to move an atom to an interstitial position, N the number of lattice sites, and N' the number of interstitial sites, one finds

$$n = (NN')^{1/2} e^{-\varepsilon/2kT} \tag{8.6}$$

by a method similar to that which gave Eq. (8.5).

Crystals of alkali halides or silver halides may contain small amounts of divalent salts such as $CaCl_2$ or $SrCl_2$. Each of their divalent metal ions can be at a lattice site that is normally occupied by a monovalent alkali ion, producing an excess of one positive ion. The defects thus formed change electrical conductivity noticeably because the positive ion excess can move in response to an electric field.

According to the elementary transport theory discussed in Chapter 1, the conductivity σ is given by

$$\sigma = ne\mu, \tag{8.7}$$

where n is the concentration of charge carriers of charge e and μ is their mobility. Thus, conductivity measurements give information on the number of charge carriers, the temperature variation of which gives the energy ε.

Problems

1. N independent particles of spin $\frac{1}{2}$ are in a constant magnetic field H. Taking the Zeeman effect into consideration, find (a) the partition function due to the spin, (b) the internal energy, (c) the total magnetic moment, (d) the heat capacity.
2. A system of charged particles with a number density $n(r)$ has a potential energy $\varphi(\mathbf{r})$. The initial energy is given by

$$U = \frac{1}{2} e^2 \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}\, d\mathbf{r}\, d\mathbf{r}' + e \int n(\mathbf{r})\varphi\, d\mathbf{r}.$$

   Assuming that the entropy of this system is given by

$$S = -k \int n[\log n(\mathbf{r}) - 1]\, d\mathbf{r},$$

   find the equation that $n(\mathbf{r})$ satisfies.
3. If n out of N molecules in a perfect crystal move from the interior of the crystal to the surface, leaving n vacant lattice sites, show that

$$\frac{n}{N+n} = e^{-w/kT},$$

   where w is the energy required to bring each of the atoms to the surface. Assume that $N \gg 1$.

4. A particle of mass m moves in the one-dimensional domain $0 \leqslant x \leqslant l$. Assuming that the collisions of the particle with the boundaries are perfectly elastic (1) derive the phase-space volume $\Gamma(E)$ corresponding to the energies of the particle less than E, and (2) show that $\Gamma(E)$ is invariant when the boundary wall $x = l$ is moved very slowly.
5. For a surface of N absorption sites, determine the absorption coefficient (the ratio of the number of absorbed atoms to the number of absorption sites) as a function of p and T under the assumption that each absorption site can take only one atom. Assume also that the absorbed atoms fall in an energy state which is below that of the atoms prior to absorption by the amount ε.
6. Show that the specific heat for a one-dimensional anharmonic oscillator of a potential $\phi(x) = ax^2 - bx^3 - cx^4$ is approximately proportional to T, the absolute temperature, and is approximately given by

$$c = k\left[1 + \left(\frac{3c}{2a^2} + \frac{15b^2}{8a^3}\right)kT\right].$$

The coefficients b and c are assumed to be small.

7. For N identical particles without interaction, find the equilibrium distribution function over states with energy ε_j and angular momentum M_j by using the constraints:

$$\sum_j \varepsilon_j N_j = U, \qquad \sum_j N_j = N, \qquad \sum_j N_j M_j = M,$$

where N_j represents the number of particles in the state j, and U, N, and M are constants.

GENERAL REFERENCE

J. O. Hirschfelder, C. R. Curtiss, and R. B. Bird, "The Molecular Theory of Gases and Liquids." Wiley, New York, 1954.

REFERENCES

1. A. Isihara, *Proc. Intern. Conf. Theoret. Phys. Kyoto and Tokyo, 1953*, p. 400.
2. F. Kroupa, ed., "Theory of Crystal Defects" Academic Press, New York, 1966.

4

Ideal Bosons and Fermions

The chapter presents basic theoretical methods for treating gases of bosons and fermions. It will be seen that descriptions based on grand ensembles are very natural and useful.

In interacting many-body systems, complicated phenomena are often treated successfully by excitons, which are independent of each other. Indeed, to find such excitons is one of the basic steps in many-body theories. If excitons do not have interactions, the methods in this chapter are applicable. We shall consider excitations in He II later in this chapter.

4.1. Blackbody Radiation

According to Kirchhoff's law, the ratio of the radiant emittance to the absorptivity of a substance in thermal equilibrium is determined only by the frequency and the temperature, independently of the substance, and is identical with that emitted by a perfect blackbody, i.e., *blackbody radiation*. The blackbody is a hypothetical object that absorbs radiation without reflection. From Kirchhoff's law, the radiation coming out of a small aperture of a furnace in thermal equilibrium—*hollow space radiation*—becomes equivalent to blackbody radiation. This can be understood by noting that photons entering a furnace through a very small hole are not likely to leave the hole immediately (Fig. 4.1), that is, the absorption of light by the hollow space is almost 100%.

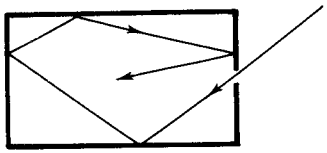


Fig. 4.1. Hollow space radiation realizes blackbody radiation.

The frequency distribution of blackbody radiation was first discovered by Planck. He explained it by introducing the notion of energy quanta and opened the door to the new world of quantum mechanics.

The radiation field in a vacuum is determined by wave equations such as

$$\nabla^2 E - \frac{1}{c^2}\frac{\partial^2 E}{\partial t^2} = 0. \tag{1.1}$$

For the radiation field in thermal equilibrium in the hollow space, we use the boundary condition that the waves are perfectly reflected at its walls. Expanding E into a Fourier series, we find that Eq. (1.1) is reduced to the equation for simple harmonic oscillators. The radiation field is equivalent to a collection of simple harmonic oscillators with a discrete energy spectrum.

According to quantum mechanics, the energies of a simple harmonic oscillator of frequency ν are given by

$$\varepsilon_\nu = (n + \tfrac{1}{2})h\nu, \qquad n = 0, 1, 2 \ldots, \tag{1.2}$$

where h is Planck's constant:

$$h = 6.62 \times 10^{-27} \text{ erg sec.}$$

The classical partition function for a one-dimensional harmonic oscillator of frequency ν is

$$Z_\nu = \frac{e^{-(1/2)h\nu/kT}}{1 - e^{-h\nu/kT}}. \tag{1.3}$$

Using Eq. (2.10) of Chapter 3, we find that the average energy is

$$u_\nu = \frac{h\nu}{2} + \frac{h\nu}{e^{h\nu/kT} - 1}. \tag{1.4}$$

At 0°K, the second term vanishes, but the first term does not. Therefore, the first term is called the *zero-point energy*.

Equation (1.4) may be used to calculate the internal energy of the radiation field. However, we must sum u_ν over all possible frequencies because u_ν is for a particular frequency ν.

For this purpose, let us consider electromagnetic waves with frequencies between ν and $\nu + d\nu$ confined in a vessel of volume V. The waves have nodes

at the boundary since they are perfectly reflected and are standing waves. Therefore, if the waves are in a rectangular solid of side lengths A, B, C in the x, y, z directions, respectively, we have the relations:

$$\frac{\lambda}{2\alpha} l = A, \qquad \frac{\lambda}{2\beta} m = B, \qquad \frac{\lambda}{2\gamma} n = C,$$

where α, β, γ are directional cosines, l, m, n, are positive integers, and λ is the wavelength. Since $\alpha^2 + \beta^2 + \gamma^2 = 1$, the integers l, m, n satisfy the following equation:

$$\left(\frac{l}{A}\right)^2 + \left(\frac{m}{B}\right)^2 + \left(\frac{n}{C}\right)^2 = \left(\frac{2}{\lambda}\right)^2. \tag{1.5}$$

The number of states of the waves of frequencies between ν and $\nu + d\nu$ can be obtained from the set of integers l, m, n satisfying the above equation, which, in turn, is given approximately by the surface of the ellipsoid represented by this equation. Since $\lambda = c/\nu$, we obtain the number by differentiating $\frac{1}{8}$ of the volume of the ellipsoid, $\frac{4}{3}\pi(V/c^3)\nu^3$, with respect to ν. Here, $V = ABC$. The number of states of frequencies less than ν is given by

$$2\frac{4\pi}{3}\frac{V}{c^3}\nu^3.$$

The factor 2 is due to the two possible directions of polarization of the waves for a given ν.

As we shall see, the same result can be obtained more quickly using Eq. (3.12) and taking into consideration the particle aspect of radiation, namely the photons. A photon of energy $h\nu$ has momentum $h\nu/c$. According to Eq. (3.12), the number of states of photons with momenta less than $h\nu/c$ is $2(4\pi/3)(h\nu/c)^3 V/h^3$. Thus, the number of states with frequencies between ν and $\nu + d\nu$ is given by $8\pi\nu^2\, d\nu(V/c^3)$.

We introduce the energy density ρ_ν defined by

$$\rho_\nu = \frac{8\pi\nu^2 u_\nu}{c^3}. \tag{1.6}$$

Then, Eq. (1.4) gives

$$\rho_\nu = \frac{4\pi h\nu^3}{c^3} + \frac{8\pi h\nu^3}{c^3}\frac{1}{e^{h\nu/kT} - 1}. \tag{1.7}$$

The first term will be divergent when summed over all frequencies. However, it is constant and may be omitted from our consideration, since radiation

corresponds not to absolute values of energy, but to differences between energy levels. Thus, expanding the denominator of Eq. (1.7) and integrating over ν, we obtain the internal energy of radiation per unit volume as follows:

$$\frac{U}{V} = \frac{8\pi h}{c^3}\int_0^\infty \frac{\nu^3\,d\nu}{e^{h\nu/kT}-1}$$

$$= \frac{8\pi h}{c^3}\int_0^\infty \nu^3\left(\sum_{n=1}^{\infty} e^{-nh\nu/kT}\right)d\nu$$

$$= \frac{48\pi h}{c^3}\left(\frac{kT}{h}\right)^4 \sum_{n=1}^{\infty}\frac{1}{n^4}. \tag{1.8}$$

Using

$$\sum_1^\infty \frac{1}{n^4} = \frac{\pi^4}{90}, \tag{1.9}$$

we obtain

$$\frac{U}{V} = \frac{8\pi^5}{15}\frac{(kT)^4}{(hc)^3}. \tag{1.10}$$

The pressure p due to radiation, i.e., the pressure of photons, is calculated by considering the changes in photon momenta upon collision with the wall. The number of photons that collide with the wall per unit time per unit volume in the solid angle $d\omega$ in the direction θ measured from the perpendicular to the wall is given by $(\rho_\nu/h\nu)c(\cos\theta)\,d\omega/4\pi V$. The momentum change is $[h\nu-(-h\nu)](\cos\theta)/c = 2h\nu(\cos\theta)/c$. Thus,

$$pV = \left[\int_0^\infty 2h\nu(\rho_\nu/h\nu)\,d\nu\right]\frac{1}{4\pi}\int_0^{\pi/2} 2\pi\cos^2\theta\sin\theta\,d\theta = \frac{1}{3}U. \tag{1.11}$$

This equation differs from Eq. (2.7) of Chapter 1 and its quantum mechanical analog, Eq. (3.19), by a factor 2. The reason is that a photon has two different directions of polarization.

Problems

1. Show that the energy density emitted per second from a unit surface of a blackbody is σT^4, where σ is given by

$$\sigma = \frac{2\pi^5 k^4}{15c^2h^3} = 5.67\times 10^{-15}\ \text{erg sec}^{-1}\,\text{cm}^{-2}\,\text{deg}^{-4}. \tag{1.12}$$

2. Prove that the wavelength λ_m, at which the energy density of blackbody radiation becomes a maximum, is determined by the equation

$$e^{-x} + \tfrac{1}{5}x - 1 = 0, \tag{1.13}$$

where $x = ch/\lambda_m kT$. The solution of this equation is approximately $x = 4.9651$, and thus

$$\lambda_m T = \frac{ch}{4.9651k} = 0.289 \quad \text{cm deg.} \tag{1.14}$$

4.2. Specific Heats of Solids

At finite temperatures, atoms or molecules forming a crystal vibrate about their equilibrium positions. These *lattice vibrations* characterize thermal properties of crystals and are very important in the theory of crystals. Generally, their treatment is called *lattice dynamics*. In 1907, Einstein (*1*) presented a simple model to treat the lattice vibrations. He assumed that the atoms in a crystal lattice vibrate with the same frequency. In 1912, Debye (*2*) improved the Einstein model by introducing a frequency range determined by the lowest possible frequencies of a continuum that has the same elastic constants as the original crystal.

The Einstein model starts with the assumption that atoms vibrating in a crystal with the same frequency ν have discrete energies given by $nh\nu$ in each direction of vibration, as in Planck's radiation theory. Each atom of the crystal may be replaced by a harmonic oscillator. Since each atom has three independent directions of vibration, the crystal lattice composed of N atoms may be considered to be equivalent to a collection of $3N$ linear harmonic oscillators. If Eqs. (1.2) and (1.3) are used, the partition function for a crystal of frequency ν may be expressed as follows:

$$Z_\nu = \left[\frac{e^{-(1/2)h\nu/kT}}{(1-e^{-h\nu/kT})}\right]^{3N}. \tag{2.1}$$

Using Eq. (2.10) of Chapter 3, we find the internal energy due to the lattice vibration as follows:

$$U = \frac{3}{2}Nh\nu + \frac{3Nh\nu}{e^{h\nu/kT}-1}. \tag{2.2}$$

We notice here that the first term is independent of temperature. It exists even at absolute zero, where the second term vanishes, and is the zero-point energy.

It is convenient to introduce the so-called *Einstein characteristic temperature* defined by

$$\frac{h\nu}{k} = \vartheta_E. \tag{2.3}$$

The heat capacity may be derived from Eq. (2.2) in the following form (see Fig. 4.2):

$$C_V = 3Nkf_E\left(\frac{\vartheta_E}{T}\right), \tag{2.4}$$

$$f_E = \frac{x^2e^x}{(e^x-1)^2}. \tag{2.5}$$

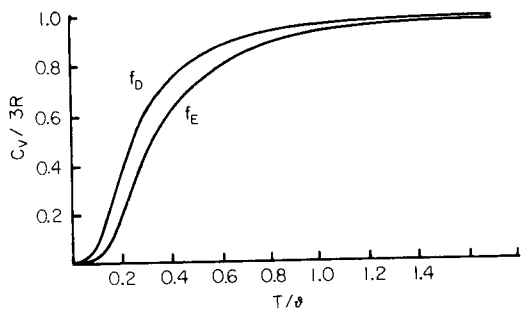


Fig. 4.2. The Debye function and the Einstein function for the specific heats of solids.

The function $f_E(x)$ approaches 1 as $x \to 0$. Thus, at high temperatures, $C_V = 3NK$, in agreement with Dulong–Petit's law. At low temperatures, the heat capacity is given by

$$C_V = 3Nk\left(\frac{\vartheta_E}{T}\right)^2 e^{-\vartheta_E/T}. \tag{2.6}$$

However, experiments show that the heat capacity varies as T^3. Thus, Einstein's theory does not agree with the data. This discrepancy is removed by Debye's improved model.

In Debye's theory, we first determine the frequencies of possible waves in a continuous elastic solid. We shall assume that the elastic body has side lengths A, B, and C in the x, y, and z directions, respectively. We assume further that the elastic waves are reflected perfectly at the boundary. Therefore, waves traveling parallel to the x axis must have wavelength $\lambda/2$ equal to A/l, where l is a positive integer. For a given type of vibration, say longitudinal, the number of waves having wavelengths larger than λ_0 is given by the

number of positive integers l, m, n satisfying

$$\left(\frac{l}{A}\right)^2 + \left(\frac{m}{B}\right)^2 + \left(\frac{n}{C}\right)^2 \leqslant \left(\frac{2}{\lambda_0}\right)^2. \tag{2.7}$$

This number is approximately given by the volume of an octant of the ellipsoid of semimajor axes A, B, C. Hence, we have $(1/8)(4\pi/3)ABC/(\lambda_0/2)^3 = (4\pi/3)(V/\lambda_0{}^3)$, where V is the volume of the box.

An elastic body has one longitudinal and two transverse vibrational modes. Therefore, the total number of independent waves with frequencies less than ν_0 is

$$\frac{4\pi}{3}V\left(\frac{2}{c_t{}^3} + \frac{1}{c_l{}^3}\right)\nu_0{}^3, \tag{2.8}$$

where c_l and c_t are the velocities of longitudinal and transverse waves, respectively.

Let us rewrite Eq. (2.8) in terms of the average velocity defined by

$$\frac{3}{c^3} = \frac{1}{c_l{}^3} + \frac{2}{c_t{}^3}. \tag{2.9}$$

The number of normal modes with frequencies between ν and $\nu + d\nu$ is given by

$$g(\nu)\,d\nu = \frac{12\pi V}{c^3}\nu^2\,d\nu. \tag{2.10}$$

Since the total number of normal modes must be equal to the total degrees of freedom, we have

$$\int_0^{\nu_D} g(\nu)\,d\nu = 3N, \tag{2.11}$$

where ν_D represents the maximum frequency. We obtain

$$g(\nu) = \frac{9N\nu^2}{\nu_D{}^3}, \qquad \nu \leqslant \nu_D,$$
$$= 0, \qquad \nu > \nu_D, \tag{2.12}$$

with

$$\nu_D = \left(\frac{3N}{4\pi V}\right)^{1/3} c. \tag{2.13}$$

It is convenient to introduce the *Debye characteristic temperature* by writing

$$\vartheta_D = \frac{h\nu_D}{k} = \frac{hc}{k}\left(\frac{3N}{4\pi V}\right)^{1/3}. \tag{2.14}$$

The distribution function may be used for evaluating the partition function. Remembering that one normal mode corresponds to one harmonic oscillator, and summing the logarithms of Z_ν given by Eq. (1.3) over the entire frequency range, we obtain the following expression for the partition function:

$$\log Z = -\frac{9}{8}\frac{Nh\nu_D}{kT} - 9N\frac{T^3}{\vartheta_D{}^3}\int_0^{\vartheta_D/T} x^2 \log(1 - e^{-x})\, dx. \tag{2.15}$$

Thus, the internal energy is

$$U = \frac{3}{2}Nh\bar{\nu} + 3NkTG\left(\frac{\vartheta_D}{T}\right), \tag{2.16}$$

where $\bar{\nu}$ is the average frequency, determined by

$$\bar{\nu} = \frac{\int_0^{\nu_D} \nu g(\nu)\, d\nu}{\int_0^{\nu_D} g(\nu)\, d\nu} = \frac{3}{4}\nu_D, \tag{2.17}$$

and

$$G(x) = \frac{3}{x^3}\int_0^x \frac{x^3\, dx}{(e^x - 1)}. \tag{2.18}$$

Differentiating Eq. (2.16) with respect to temperature T and using the equality

$$\frac{dG(x)}{dx} = -\frac{3}{x}G(x) + \frac{3}{e^x - 1}, \tag{2.19}$$

one finds the following expression for the heat capacity C_V:

$$C_V = 3Nkf_D\left(\frac{\vartheta_D}{T}\right), \tag{2.20}$$

where $f_D(x)$ is the Debye function defined by (see Fig. 4.2)

$$f_D(x) = \frac{3}{x^3}\int_0^x \frac{e^x x^4\, dx}{(e^x - 1)^2}. \tag{2.21}$$

In terms of $G(x)$, $f_D(x)$ is given by

$$f_D(x) = 4G(x) - \frac{3x}{e^x - 1}. \tag{2.22}$$

At high temperatures, the Debye function approaches 1 because

$$\lim_{x\to 0} G(x) = 1, \qquad \lim_{x\to 0}\frac{x}{e^x - 1} = 1. \tag{2.23}$$

Again we obtain Dulong–Petit's value $3Nk$ in this limit. At low temperatures, we observe that

$$\begin{aligned}\lim_{x\to\infty}\frac{x^3}{3}G(x) &= \int_0^\infty \frac{x^3\, dx}{e^x - 1} \\ &= \int_0^\infty x^3 e^{-x}(1 + e^{-x} + e^{-2x} + \cdots)\, dx \\ &= \sum_{n=1}^{\infty}\frac{6}{n^4} = \frac{\pi^4}{15};\end{aligned} \tag{2.24}$$

$$\lim_{x\to\infty}\frac{3x}{e^x - 1} = 0.$$

Thus, we obtain

$$C_V = Nk\frac{12}{5}\pi^4\left(\frac{T}{\vartheta_D}\right)^3 = 464.4\left(\frac{T}{\vartheta_D}\right)^3 \text{ cal deg}^{-1}\text{ mole}^{-1}. \tag{2.25}$$

According to the Debye formula, the heat capacity is determined entirely by $f_D(\vartheta_D/T)$. Therefore, a law of corresponding states holds if C_V is plotted against the reduced temperature ϑ_D/T. This has been confirmed by experiment. In Fig. 4.2, we compare $f_D(T/\vartheta_D)$ with $f_E(T/\vartheta_E)$.

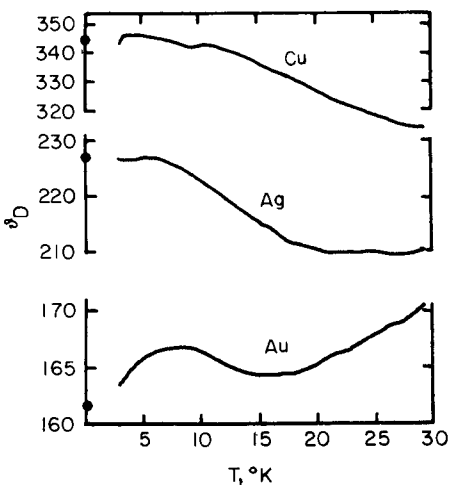


Fig. 4.3. Debye temperature as a function of the temperature. The black points on the ordinate represent ϑ from 0°K elastic constant. (For LiF, see Fig. 8.10.) [After Martin (*3*).]

Using Eqs. (2.14) and (2.20), we can determine ϑ_D in two ways, by measuring the velocities of elastic waves and by measuring the specific heats. The two methods give fairly good agreement. This can be considered as experimental justification of Debye's continuum assumption.

However, it has been found that ϑ_D determined at various temperatures to give the best fit of heat capacity depends on the temperatures instead of being constant (*3*) (Fig. 4.3). This fact indicates that the distribution function used by Debye is not perfect. Therefore, efforts have been made to obtain an improved distribution function using a more realistic lattice model instead of the continuum model. We shall discuss this subject in a later section.

4.3. Quantum Statistics of Ideal Gases

The energy expressions (1.4) and (2.2) have a very important feature. They are characterized by $e^{h\nu/kT} - 1$ in the denominator, which yields the classical energy expression kT at high temperatures. The deviations of the energy expressions from kT become larger at lower temperatures. The deviations of energy from the classical value depend on Planck's constant h.

Planck's efforts to explain the frequency distribution of radiation, and Einstein's and Debye's theories of the heat capacities of solids, initiated the development of quantum mechanics and quantum statistical mechanics. As we shall discuss in this section, the form of Eqs. (1.4) and (2.2) can be understood more generally in terms of quantum statistical mechanics.

We discussed in Section 3.3 the necessity of modifying the classical partition function to meet the quantum mechanical uncertainty principle and the indistinguishability of particles. However, the Pauli principle of quantum mechanics requires us to construct quantum statistical mechanics. This principle states that the wave function describing a system of identical particles must be either symmetric or antisymmetric upon exchange of the particles.

The particles having symmetric wave functions are called *bosons*. It is known that any number of bosons can be in one energy level. In this case, we say that the particles obey *Bose–Einstein statistics*. Photons and atoms composed of an even number of elementary particles are *bosons*. Particles having antisymmetric wave functions are called *fermions*. No more than one fermion can be in the same energy state. We say that these particles obey *Fermi–Dirac statistics*. Electrons and atoms composed of odd numbers of fermions obey Fermi–Dirac statistics.

Let us derive the partition functions of ideal Bose and Fermi gases. Consider a system with N_i bosons or fermions in the energy level ε_i, with N_i

being 0 or 1 for fermions and arbitrary for bosons. The total energy is

$$\sum_i N_i \varepsilon_i = E, \tag{3.1}$$

and the total number of particles is given by

$$\sum_i N_i = N. \tag{3.2}$$

By definition, the partition function is

$$Z_N = \sum_{\{N_i\}} \exp\left(-\sum_i \frac{N_i \varepsilon_i}{kT}\right), \tag{3.3}$$

where $\{\varepsilon_i\}$ represents the energy levels of the total system of N particles. The distribution of particles is characterized by the set of integers N_i. The summation is subject to the restricting condition of Eq. (3.2). The partition function used here is not expressed as a product of the partition functions for one particle as in Eq. (4.14) of Chapter 2 because the specification of the number of particles in respective energy levels is important for our present discussion.

The grand partition function assumes the form

$$\begin{aligned} \Xi_0 &= \sum_N \lambda^N Z_N \\ &= \sum_{\{N_i\}} \prod \left(\lambda \exp \frac{-\varepsilon_i}{kT}\right)^{N_i}. \end{aligned} \tag{3.4}$$

On the right-hand side of the first equation, the summation over N is extended from zero to infinity, enabling us to drop the restricting condition of Eq. (3.2). In other words, N_i takes on all values between zero and infinity for bosons and either 0 or 1 for fermions. Equation (3.4) can be transformed into

$$\Xi_0 = \prod_i \sum_{N_i} \left(\lambda \exp \frac{-\varepsilon_i}{kT}\right)^{N_i}. \tag{3.5}$$

On the right-hand side of Eq. (3.5) we perform the summation over all possible N_i first for a given energy level and then multiply the terms corresponding to various energy levels. Equation (3.5) may be proved by applying Cauchy's integral theorem to the right-hand side of Eq. (3.4) and then summing over N prior to taking the product, namely

$$\begin{aligned} \Xi_0 &= \sum_N \lambda^N \prod_i \frac{1}{2\pi i} \oint x^{-(N+1)} \sum_{N_i} [e^{-\varepsilon_i/kT} x]^{N_i} \, dx \\ &= \prod_i \frac{1}{2\pi i} \oint (x - \lambda)^{-1} \sum_{N_i} [e^{-\varepsilon_i/kT} x]^{N_i} \, dx \\ &= \prod_i \sum_{N_i} [\lambda e^{-\varepsilon_i/kT}]^{N_i}. \end{aligned} \tag{3.6}$$

The summation on the right-hand side of Eq. (3.6) extends to infinity for bosons and to one for fermions. We finally obtain

$$\Xi_0 = \prod_i [1 \mp \lambda e^{-\varepsilon_i/kT}]^{\mp 1}, \tag{3.7}$$

where the upper and the lower signs correspond, respectively, to bosons and fermions.

Using Eq. (4.11) of Chapter 3, the total number of particles is given by

$$N = \sum_i \frac{1}{\lambda^{-1} e^{\varepsilon_i/kT} \pm 1}. \tag{3.8}$$

Each term on the right-hand side of this equation represents the average number of particles in the respective energy levels and is called the *Bose–Einstein distribution* (minus sign) or the *Fermi–Dirac distribution* (plus sign). If we denote these distribution functions by $f^{(0)}(\varepsilon)$, we have

$$\begin{aligned} f^{(0)}(\varepsilon) &= \frac{1}{\lambda^{-1} e^{\varepsilon/kT} - 1} \qquad \text{(bosons)}, \\ f^{(0)}(\varepsilon) &= \frac{1}{\lambda^{-1} e^{\varepsilon/kT} + 1} \qquad \text{(fermions)}. \end{aligned} \tag{3.9}$$

In these expressions, $\lambda = e^{\mu/kT}$ is to be determined as a function of N and T from Eq. (3.8). If, as in the case of radiation, the total number N is not specified, μ should be zero, and $\lambda = 1$.

Let us consider ideal bosons or fermions in a cubic vessel of side length L. According to quantum mechanics, their energy takes on only discrete values given by

$$\varepsilon_i = \frac{h^2}{8mL^2}(n_1^2 + n_2^2 + n_3^2), \tag{3.10}$$

where n_i, $i = 1, 2, 3$, are positive integers and m is the mass of a particle. A given set (n_1, n_2, n_3) determines one energy state. If L is sufficiently large, the energy levels will be distributed almost continuously. Then, the number of states below a certain value ε of energy will be given by $\frac{1}{8}$ of the volume of the sphere of radius $(8mL^2\varepsilon/h^2)^{1/2}$ and with origin at $n_1 = n_2 = n_3 = 0$. This is $(8mL^2\varepsilon/h^2)^{3/2}\pi/6$. Thus, the number of states with energies between ε and $\varepsilon + d\varepsilon$ is given by

$$g(\varepsilon)\, d\varepsilon = 2\pi\left(\frac{2m}{h^2}\right)^{3/2} V\varepsilon^{1/2}\, d\varepsilon, \tag{3.11}$$

where V is the volume of the system.

The term $g(\varepsilon)$ is called the *density of states*. If the particles have spin $\frac{1}{2}$, each energy state may allow two particles of opposite spins. In such cases, $g(\varepsilon)$ must be multiplied by a factor of two. For an ideal gas particle, $\varepsilon = p^2/2m$ and in the absence of spin

$$g(\varepsilon)\, d\varepsilon = V\frac{4\pi p^2\, dp}{h^3}, \tag{3.12}$$

in agreement with the number of states for radiation obtained in Section 4.1.

Using Eq. (4.11) of Chapter 3 and replacing summation by integration, one finds the equation of state as follows:

$$\frac{pV}{kT} = \mp \int_0^\infty \log(1 \mp \lambda e^{-\varepsilon/kT}) g(\varepsilon)\, d\varepsilon, \tag{3.13}$$

where the upper and lower signs correspond, respectively, to the Bose and Fermi gas. The absolute activity λ is determined by the condition (4.11) of Chapter 3 that the total number of particles is N. Using a Taylor expansion and integrating term by term, we obtain from Eq. (3.13)

$$\frac{p}{kT} = \frac{(2\pi mkT)^{3/2}}{h^3} G_{5/2}(\lambda), \tag{3.14}$$

and from Eq. (4.11) of Chapter 3,

$$n = \frac{(2\pi mkT)^{3/2}}{h^3} G_{3/2}(\lambda); \qquad n = \frac{N}{V}; \tag{3.15}$$

where

$$G_s(\lambda) = \sum_{l=1}^{\infty} \frac{(\pm 1)^{l+1}\lambda^l}{l^s}. \tag{3.16}$$

The plus sign represents the Bose function and the minus sign the Fermi function.

Eliminating λ from Eqs. (3.14) and (3.15), we can express p as a function of density. At high temperatures, we find

$$\begin{aligned} \frac{pv}{kT} &= 1 \mp \frac{1}{2^{5/2}} \frac{h^3 n}{(2\pi mkT)^{3/2}} + \cdots \\ &= 1 \mp \frac{1}{2^{5/2}} \lambda_0 + \cdots, \quad (- \text{Bose}; + \text{Fermi}) \end{aligned} \tag{3.17}$$

where λ_0^{-1} is the partition function per particle for specific volume v of a classical gas [cf. Eq. (5.19) of Chapter 3]. Here, the second and higher terms

are due to quantum statistics. The effect of quantum statistics is large for small masses, low temperatures, and high densities. The series expansion on the right-hand side of this equation is valid if $\lambda_0 < 1$ or

$$n < \frac{(2\pi mkT)^{3/2}}{h^3}. \tag{3.18}$$

The quantity $h/(2\pi mkT)^{1/2}$ is called the thermal de Broglie wavelength; $h/(3mkT)^{1/2}$ corresponds to the thermal energy $3kT/2$ of a particle. Quantum corrections become necessary when the mean particle separation of our system becomes of the order of the de Broglie wavelength. When $n \geqslant (2\pi mkT)^{3/2}/h^3$, the gas is *degenerate*. Hydrogen and helium gases are degenerate at low temperatures, whereas an electron gas is degenerate even at room temperature.

Equation (3.17) is a high-temperature expansion. It is not easy to evaluate the equation of state at arbitrary temperatures. However, we can show that the Bernoulli equation

$$pV = \tfrac{2}{3}U \tag{3.19}$$

holds also for ideal quantum gases. According to Eq. (3.15), λ is a function of $vT^{3/2}$, and therefore $N\mu/T$ is a function of $vT^{3/2}$. Furthermore, Eq. (3.14) indicates that pV/T is dependent on $vT^{3/2}$. Thus, $F/T = (N\mu - pV)/T$ is a function of $vT^{3/2}$. Further use of Eqs. (2.12) and (2.14) of Chapter 3 gives the above expression.

Problem

Determine the ratio of the specific heat at constant volume to that at constant pressure for an ideal quantum gas at high temperatures.

4.4. Bose–Einstein Condensation

In the previous section, both Fermi–Dirac and Bose–Einstein gases were treated together. At low temperatures, it is no longer possible to treat them at the same time, because of their very different behaviors. We shall first discuss an ideal Bose gas (*4*).

According to the Bose–Einstein distribution given by Eq. (3.9), the number of particles in the lowest energy state $\varepsilon = 0$ is

$$N_0 = \frac{1}{\lambda^{-1} - 1}. \tag{4.1}$$

Since N_0 has to be positive, λ cannot be larger than 1. If λ is not close to unity, the density $n_0 = N_0/V$ of such particles is small. On the other hand,

if λ is close to unity, N_0 becomes of the order of N. If we replace the summation in Eq. (3.8) by integration using Eq. (3.11), we shall lose N_0, the contribution from $\varepsilon = 0$ state, because Eq. (3.11) gives a weight of zero to this state. If λ is close to unity, N_0 is important, and the term $\varepsilon = 0$ has to be taken into consideration separately in Eq. (3.8). The other terms may be summed by using integration. It is a sufficiently accurate approximation to use $\varepsilon = 0$ as the lower limit. We arrive at

$$n = n_0 + \frac{(2\pi mkT)^{3/2}}{h^3} G_{3/2}(\lambda), \tag{4.2}$$

where, as in Eq. (3.16),

$$G_s(\lambda) = \sum_{l=1}^{\infty} \frac{\lambda^l}{l^s} \tag{4.3}$$

$G_{3/2}(\lambda)$ is an increasing function of its argument. Its maximum occurs at $\lambda = 1$, namely $G_{3/2}(1) = 2.612$. Let us introduce a temperature T_0 defined by

$$n = \frac{(2\pi mkT_0)^{3/2}}{h^3} G_{3/2}(1). \tag{4.4}$$

From Eqs. (4.2) and (4.4), we see that $(n - n_0)/n < 1$ for $T < T_0$. That is, the particles in the lowest energy state give rise to a finite contribution to the total density. Using the fact that $G_{3/2}(\lambda)$ varies slowly with λ, one obtains the following approximate results:

$$\begin{aligned} &n_1 = n\left(\frac{T}{T_0}\right)^{3/2}, \quad && n_0 = n\left[1 - \left(\frac{T}{T_0}\right)^{3/2}\right], && T < T_0 \\ &n_1 = n, && n_0 = 0, && T > T_0 \end{aligned} \tag{4.5}$$

where n_1 is the density of the particles in the excited states. The density n_0 increases as the temperature is lowered from T_0 and reaches the total density n at 0°K (Fig. 4.4). Correspondingly, n_1 decreases from n at T_0 with the temperature dependence $T^{3/2}$.

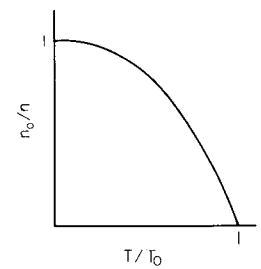


Fig. 4.4. The number density of the particles in the lowest state of an ideal Bose gas.

These theoretical results do not agree quantitatively with experimental data on liquid helium. Figure 4.5 shows the temperature variation of n_1/n

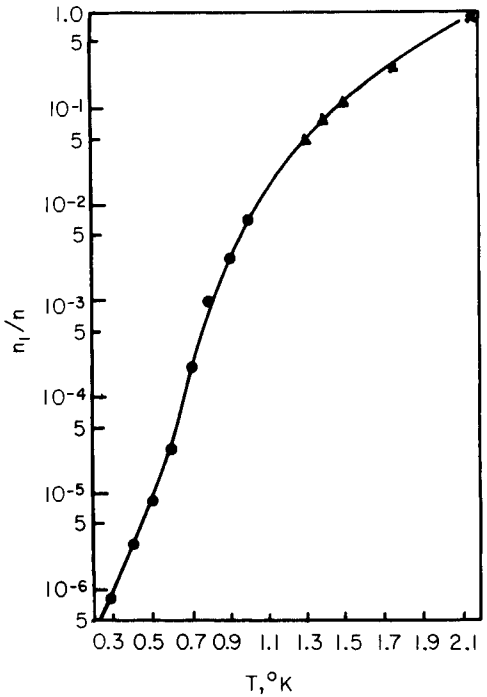


Fig. 4.5. The ratio n_1/n determined from experiments for liquid helium. [After Yarnell *et al.* (*5*).]

determined from experiments by Yarnell *et al.* (*5*). As can be seen, the temperature dependence is much stronger than $T^{3/2}$. It is close to $T^{5.5}$.

The sudden onset of an increase of n_0 at T_0 gives rise to discontinuities in thermodynamic functions of the perfect gas. For instance, as the temperature falls below T_0, the pressure becomes approximately independent of volume. This is so because, below T_0, $G_s(\lambda)$ may be replaced by $G_s(1)$, and Eq. (3.14) gives

$$\frac{p}{kT} = \frac{(2\pi mkT)^{3/2}}{h^3} G_{5/2}(1) = n_1 \frac{G_{5/2}(1)}{G_{3/2}(1)}. \tag{4.6}$$

Accordingly, the pressure is a function of temperature alone and is independent of the volume. For $T > T_0$, n_0 is negligible and the pressure is determined by Eqs. (3.14) and (3.15) as before. The pressure p given by Eq. (4.6) is less than that obtained from Eq. (3.14).

The internal energy may be found from Eq. (3.19). The heat capacity may then be calculated by differentiating the internal energy with respect to T. We find that, at $T = T_0$, a discontinuity in the slope of the heat capacity appears.

Because of these differences in the properties above and below T_0, we

conclude that a phase transition takes place at T_0. This transition is caused by the condensation of particles into the lowest state due to Bose–Einstein statistics and is called *Bose–Einstein condensation*, or simply Bose condensation.

We have treated an ideal Bose gas. However, the condensation resembles the λ-transition in liquid helium. Ordinary helium atoms are composed of two neutrons and two protons, and obey Bose–Einstein statistics. The λ-transition takes place in liquid helium at 2.18°K (λ-point), which is very close to T_0 calculated from Eq. (4.4), namely 3.13°K. At the λ-point, the heat capacity curve shows a discontinuity that looks like the Greek letter λ.

There are still other differences between the λ-transition and the Bose condensation in an ideal gas. The heat capacity of liquid helium is logarithmically divergent at the λ-point, and the phase change is in the second order as compared with the third order in an ideal Bose gas. Nevertheless, the important role played by Bose statistics is clear because liquid ^{3}He, obeying Fermi–Dirac statistics, does not show such a transition.

In Fig. 4.6, the internal energy U, the pressure p, and the heat capacity C_V of an ideal Bose gas are schematically illustrated.

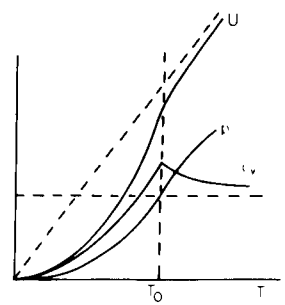


Fig. 4.6. Schematic curves showing the properties of an ideal Bose gas.

The function $G_s(\lambda)$ appears very often when we treat a Bose gas. Let us discuss its mathematical properties near the critical temperature. For this purpose, let us define (*6*)

$$\lambda = e^{-\alpha}, \tag{4.7}$$

and apply the Mellin transform. We find

$$\begin{aligned} \mathcal{G}_s(q) &= \int_0^\infty \sum_{n=1} n^{-s} e^{-n\alpha} \alpha^{q-1}\, d\alpha \\ &= \Gamma(q)\zeta(q+s), \end{aligned} \tag{4.8}$$

where $\zeta(q+s)$ is the Riemann zeta-function. The inverse transform is

$$G_s(\alpha) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \mathcal{G}_s(q) \alpha^{-q}\, dq, \tag{4.9}$$

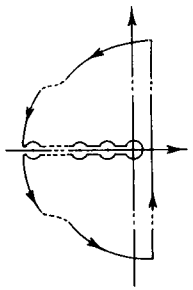


Fig. 4.7. The contour used for the Mellin transform of Eq. (4.9).

where the contour may be chosen as in Fig. 4.7, bypassing the poles of the integrand. Since $\zeta(q+s)$ has a simple pole at $s=1-q$ with a residue 1, and $\Gamma(s)$ has poles at $s=-n$ with residues $(-)^n/n!$, we obtain

$$G_s(\alpha)=\Gamma(1-s)\alpha^{s-1}+\sum_{n=0}^{\infty}\frac{(-)^n}{n!}\zeta(s-n)\alpha^n. \tag{4.10}$$

Differentiating this result, we obtain

$$\frac{\partial^n G_s(\alpha)}{\partial\alpha^n}=(-)^n G_{s-n}(\alpha). \tag{4.11}$$

Where $\alpha\to 0$, $G_s(\alpha)$ diverges as $\alpha^{-|s-1|}$ for $s<1$ and as $\log(1/\alpha)$ for $s=1$. It converges to $\zeta(s)$ if $s>1$. For $1<s\leqslant 2$, $G_s(\alpha)$ has an infinite slope at the origin, and a branch cut must be made from $\alpha=0$ to $\alpha=-\infty$.

The series (4.10) converges rapidly in the neighborhood of $\alpha=0$ for $s>0$ if s is not too large. Thus, we obtain the expansions for $\alpha\leqslant 1$

$$\begin{aligned}
G_{1/2}(\alpha)&=1.77\alpha^{-1/2}-1.46+0.208\alpha-0.0128\alpha^2,\\
G_{3/2}(\alpha)&=-3.54\alpha^{1/2}+2.61+1.46\alpha-0.104\alpha^2+0.00425\alpha^3,\\
G_{5/2}(\alpha)&=2.36\alpha^{3/2}+1.34-2.61\alpha-0.730\alpha^2+0.0347\alpha^3.
\end{aligned} \tag{4.12}$$

These expansions are useful to find the properties of a Bose gas near the condensation point, which corresponds to $\alpha=0$. In some cases, it is also useful to adopt the integral form

$$\begin{aligned}
G_s(\alpha)&=\frac{1}{\Gamma(s)}\int_0^{\infty}\frac{x^{s-1}}{e^{x+\alpha}-1}\,dx\\
&=\frac{1}{\Gamma(s)}\left[\int_1^{\infty}\frac{1}{x-e^{-\alpha}}-\frac{1}{x}(\log x)^{s-1}\,dx\right].
\end{aligned} \tag{4.13}$$

Figure 4.8 illustrates the Bose functions for $1\geqslant\alpha\geqslant 0$.

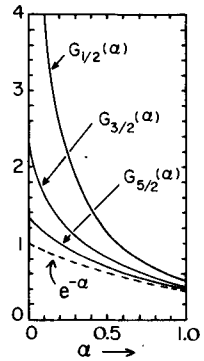


Fig. 4.8. The Bose functions $G_s(\alpha)$ for $1>\alpha>0$.

4.5. Phonons and Rotons

In 1941, Landau (*7*) attributed the peculiar properties of He II to its energy spectrum. Later, in 1947 (*8*), he introduced a refined energy spectrum to explain Peshkov's experiment (*9*) on second sound in the liquid.

From the facts that He II is not a gas and that the specific heat varies as T^3 at low temperatures, Landau conjectured the existence of phonon excitations. The phonon part of the energy spectrum is given by

$$\varepsilon(p)=cp, \qquad p\ll p_0, \tag{5.1}$$

where c is the velocity of sound. At higher temperatures, rotational motions of small clusters of helium atoms may be possible. Landau introduced a small excitation energy for such motions and a particle-like energy spectrum near a certain momentum p_0:

$$\varepsilon(p)=\Delta+\frac{(p-p_0)^2}{2m_{\text{eff}}}, \qquad p\sim p_0, \tag{5.2}$$

where m_{eff} is the effective mass of the excitons, called *rotons*.

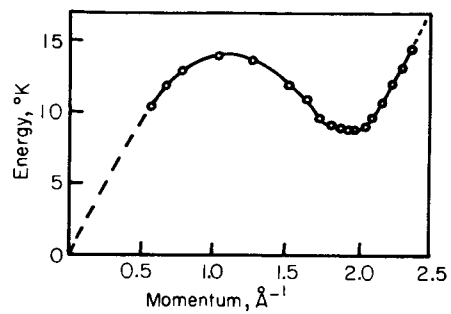


Fig. 4.9. The energy spectrum of liquid helium at 1.1°K. [After Yarnell *et al.* (*5*).]

These two parts of the energy spectrum are actually connected. The energy spectrum has been determined by neutron diffraction and other methods. Figure 4.9 shows the direct determination carried out at 1.1°K of the spectrum by inelastic neutron scattering. The three constants in Eq. (5.2) are found as

$$\begin{aligned} \frac{p_0}{\hbar} &= (1.92 \pm 0.01) \quad \text{Å}^{-1}, \\ \frac{\Delta}{k} &= (8.65 \pm 0.04) \quad \text{°K}, \\ m_{\text{eff}} &= (0.16 \pm 0.01)m, \end{aligned} \tag{5.3}$$

where m is the mass of a helium atom. The energy gap Δ is slightly temperature-dependent due to the interactions between the excitons. The following empirical relation has been proposed:

$$\frac{\Delta}{k} = 8.68 - 0.0084T^7 \quad \text{°K}. \tag{5.4}$$

The thermodynamic properties of He II are determined by the energy spectrum of elementary excitations. Since phonons and rotons can be considered as obeying Bose statistics, their numbers are given by

$$\begin{aligned} N_{\text{p}} &= \frac{4\pi V}{h^3} \int \frac{p^2\,dp}{e^{cp/kT} - 1}, \\ N_{\text{r}} &= \frac{4\pi V}{h^3} \int \frac{p^2\,dp}{\exp\{(\Delta/kT) + [(p - p_0)^2/2m_{\text{eff}}kT]\} - 1}. \end{aligned} \tag{5.5}$$

In the denominators of Eqs. (5.5), no absolute activity enters. (See the comment on p. 82.)

At low temperatures, phonons of small momentum yield major contributions to the thermodynamic properties. Thus, we can extend the integral for the phonon number to infinity:

$$\begin{aligned} N_{\text{p}} &= 4\pi\left(\frac{kT}{hc}\right)^3 V \int_0^\infty \frac{x^2\,dx}{e^x - 1} \\ &= \frac{V}{\pi^2}\left(\frac{kT}{\hbar c}\right)^3 \zeta(3), \end{aligned} \tag{5.6}$$

where $\zeta(3) = 1.202$. Using $\zeta(4) = \pi^4/90$, we find the phonon contribution to the free energy to be

$$\begin{aligned} F_{\text{p}} &= 4\pi kT\left(\frac{kT}{hc}\right)^3 V \int_0^\infty \log(1 - e^{-x})x^2\,dx \\ &= -\frac{\pi^4 kT}{90\zeta(3)} N_{\text{p}}, \end{aligned} \tag{5.7}$$

and the heat capacity is given by

$$C_{V_{\text{p}}} = \frac{2\pi^4 k}{15\zeta(3)} N_{\text{p}}. \tag{5.8}$$

To find the roton contribution, we make use of $\Delta \gg kT$. Approximating the integral for N_{r}, we obtain

$$N_{\text{r}} = \frac{4\pi V}{h^3}\left(\exp\frac{-\Delta}{kT}\right)\int\left[\exp\frac{-(p - p_0)^2}{2m_{\text{eff}}kT}\right]p^2\,dp. \tag{5.9}$$

Because the major contribution of the p-integration comes from the region $p \sim p_0$, the interval of integration may again be extended to infinity. We obtain an approximate result

$$N_{\text{r}} = \frac{2(m_{\text{eff}}kT)^{1/2}p_0{}^2 V}{(2\pi)^{3/2}\hbar^3} e^{-\Delta/kT}. \tag{5.10}$$

In terms of N_{r}, the roton contribution to the free energy is given by

$$F_{\text{r}} = -kTN_{\text{r}} \tag{5.11}$$

and the heat capacity is

$$C_{V_{\text{r}}} = kN_{\text{r}}\left[\frac{3}{4} + \frac{\Delta}{kT} + \left(\frac{\Delta}{kT}\right)^2\right]. \tag{5.12}$$

These results have been found to agree with experiment.

4.6. Heat Capacities of Fermi Gases and Fermi Liquids

The distribution function of an ideal Fermi gas is

$$f^{(0)}(\varepsilon) = \frac{1}{e^{(\varepsilon - \mu)/kT} + 1}, \tag{6.1}$$

as in Eq. (3.9). At 0°K, this distribution function assumes the values:

$$f^{(0)}(\varepsilon) = \begin{cases} 1, & \varepsilon < \varepsilon_0 \\ 0, & \varepsilon > \varepsilon_0. \end{cases} \tag{6.2}$$

where ε_0 is the chemical potential at 0°K. Using Eq. (3.8), we find

$$\varepsilon_0 = \frac{h^2}{2m}\left(\frac{3n}{4\pi g}\right)^{2/3}, \tag{6.3}$$

where g is a degeneracy factor due to the spin. For electrons, $g = 2$, and for spinless particles, we can choose $g = 1$. For monovalent metals, ε_0 is approximately 10 eV, or equivalent to 11,606°K. Instead of ε_0, one frequently uses k_0 defined by

$$\varepsilon_0 = \frac{\hbar^2}{2m}k_0^2, \qquad k_0 = 2\pi\left(\frac{3N}{4\pi V g}\right)^{1/3}; \tag{6.4}$$

where k_0 is the radius of the *Fermi sphere* that represents Eq. (6.2) in k-space.

The internal energy at 0°K, known as the zero-point energy, is

$$U_0 = \tfrac{3}{5}N\varepsilon_0. \tag{6.5}$$

The distribution function given by Eq. (6.2) and the zero-point energy are due to the Pauli exclusion principle, which allows no more than one fermion in the same energy state. Using Eq. (6.5) and the general form of Eq. (3.19), we obtain the *zero-point pressure*

$$p_z = \tfrac{2}{5}n\varepsilon_0 = cV^{-5/3}, \tag{6.6}$$

where c is a proportionality constant.

As the temperature rises from 0°K, the Fermi distribution function changes its steplike shape. As shown in Fig. 4.10, the particles corresponding

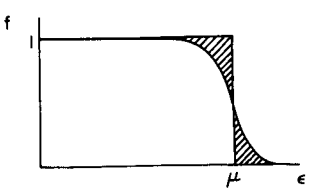


Fig. 4.10. The Fermi distribution at low temperatures.

to the shaded area with energies less than μ are excited above $\mu \sim \varepsilon_0$. By an elementary estimation of the shaded area, we find that approximately NkT/ε_0 particles are in the states above ε_0. Since the excitation energy is of the order of kT, we anticipate that the internal energy is proportional to $(kT)^2$.

An improved method to confirm this fact was given by Sommerfeld (*10*). Let us start with the integrals for the internal energy and the number of fermions:

$$U = \frac{3}{2}N\left(\frac{1}{\varepsilon_0}\right)^{3/2}\int_0^\infty \varepsilon^{3/2}f^{(0)}(\varepsilon)\,d\varepsilon, \tag{6.7}$$

$$N = \frac{3}{2}\left(\frac{1}{\varepsilon_0}\right)^{3/2}N\int_0^\infty \varepsilon^{1/2}f^{(0)}(\varepsilon)\,d\varepsilon. \tag{6.8}$$

These integrals have the form

$$J = \int_0^\infty \phi(\varepsilon)f^{(0)}(\varepsilon)\,d\varepsilon, \tag{6.9}$$

where $\phi(\varepsilon)$ is considered to be a slowly varying function of ε.

We note that $df^{(0)}/d\varepsilon$ has a sharp maximum at $\varepsilon = \mu$, where

$$-\left[\frac{\partial(\log f^{(0)}}{\partial(\log \varepsilon)}\right]_{\varepsilon=\mu} = \frac{1}{2}\frac{\mu}{kT} \gg 1. \tag{6.10}$$

For low temperatures, near 0°K, μ is close to ε_0.

Integrating Eq. (6.9) by parts, we find

$$J = -kT\int_{-\infty}^{\infty}\phi_1(x)\frac{df^{(0)}(x)}{dx}\,dx, \tag{6.11}$$

where

$$\phi_1(x) = \int_0^x \phi(x)\,dx, \qquad x = \frac{\varepsilon - \mu}{kT}, \tag{6.12}$$

and where the lower boundary $-\mu/kT$ is replaced by $-\infty$.

We expand $\phi_1(x)$ about 0:

$$\phi_1(x) = \phi_1(0) + x\left(\frac{d}{dx}\phi_1\right)_0 + \frac{1}{2}x^2\left(\frac{d^2}{dx^2}\phi_1\right)_0 + \cdots \tag{6.13}$$

and use the formulas

$$\begin{gathered}\int_{-\infty}^{\infty}\frac{x^{2m}}{(1+e^x)(1+e^{-x})}\,dx = 2(2m!)\sum\frac{(-1)^{j+1}}{j^{2m}},\\ \sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j^2} = \frac{\pi^2}{12}, \qquad \sum_{j=1}^{\infty}\frac{(-1)^{j+1}}{j^4} = \frac{7\pi^4}{720},\end{gathered} \tag{6.14}$$

to obtain

$$J = \int_0^\mu \phi(\varepsilon)\,d\varepsilon + \frac{\pi^2}{6}(kT)^2\left(\frac{d^2\phi}{d\varepsilon^2}\right)_{\varepsilon=\mu} + \frac{7\pi^4}{360}(kT)^4\left(\frac{d^4\phi}{d\varepsilon^4}\right)_{\varepsilon=\mu} + \cdots. \tag{6.15}$$

Using this result, we easily obtain the following approximate expressions:

$$N = N\left(\frac{\mu}{\varepsilon_0}\right)^{3/2}\left[1 + \frac{\pi^2}{8}\left(\frac{kT}{\varepsilon_0}\right)^2\right],$$

$$U = \frac{3}{5}N\varepsilon_0\left[1 + \frac{5}{12}\pi^2\left(\frac{kT}{\varepsilon_0}\right)^2\right], \qquad (6.16)$$

$$C_V = N\frac{\pi^2}{2}\frac{k^2T}{\varepsilon_0}\left[1 - \frac{3}{10}\pi^2\left(\frac{kT}{\varepsilon_0}\right)^2\right],$$

$$p = \frac{2}{5}n\varepsilon_0\left[1 + \frac{5}{12}\pi^2\left(\frac{kT}{\varepsilon_0}\right)^2\right].$$

These results are shown schematically in Fig. 4.11.

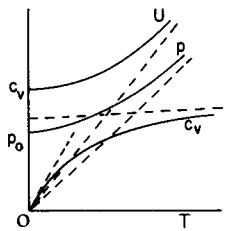


Fig. 4.11. Schematic curves for the properties of an ideal Fermi gas.

From these expressions, we learn that the internal energy is proportional to $(kT)^2$, that the heat capacity varies as k^2T/ε_0, and that the pressure does not vanish at absolute zero. As defined in Eq. (6.6), $(2/5)n\varepsilon_0$ is the zero-point pressure. For monovalent metals, the zero-point pressure is of the order of 200,000 atm. This large pressure is, however, compensated by the attractive forces due to the ions, and the electrons do not escape from the metals.

The linear dependence of the heat capacity on temperature is one of the most important results of quantum statistics. The experimental proportionality constant γ in the expression

$$\frac{C_V}{N} = c_V = \gamma T \qquad \text{(per atom)}, \qquad (6.17)$$

is slightly larger than the theoretical predictions:

$$\gamma = \frac{\pi^2}{2}\frac{k^2}{\varepsilon_0}z. \qquad (6.18)$$

Here, z is the number of valence electrons per atom and k is the Boltzmann constant. For 1 mole of a monovalent metal, for which $z = 1$, the proportionality constant is

$$\gamma_m = 3.25 \times 10^{-5}V_m^{2/3}n^{1/3} \quad \text{cal mole}^{-1}\,\text{deg}^{-1}, \qquad (6.19)$$

where V_m is the molar volume and n is the number of moles.

Figure 4.12 illustrates the specific heat of pure gold measured by Martin

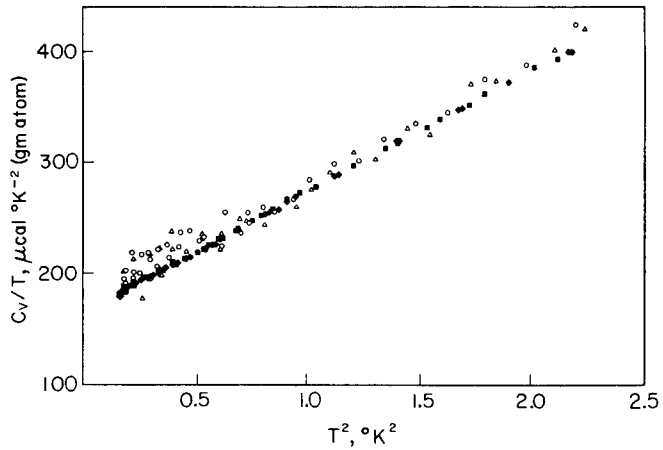


Fig. 4.12. Experimental electronic specific heat for gold. [After Martin (*3*).]

(*3*). Since the lattice contribution is proportional to T^3, the specific heat divided by T is plotted against T^2. The result is a straight line, its intersection with the ordinate giving γ. His results on γ [in μcal °K^{-2} (gm atom)] and Debye constants at 0°K are

Cu: $\gamma = 165.2 \pm 0.8$, $\vartheta = 345.8 \pm 1.2$,
Ag: $\gamma = 153.1 \pm 0.9$, $\vartheta = 227.3 \pm 0.6$,
Au: $\gamma = 165.1 \pm 1.0$, $\vartheta = 162.3 \pm 0.5$.

In Table 4.1, we give the ratio of the experimental and theoretical values of γ. It should be noted that the experimental γ values are generally higher than the theoretical prediction. For sodium, γ is about 20% higher than the free-electron value.

TABLE 4.1

Constants for Electronic Heat Capacities

	Cu	Ag	Au	Li	Na	K
ε_0 (eV)	7.04	5.51	5.51	4.76	3.20	2.12
γ/γ_{free}	1.38	1.01	1.09	2.32	1.22	1.10

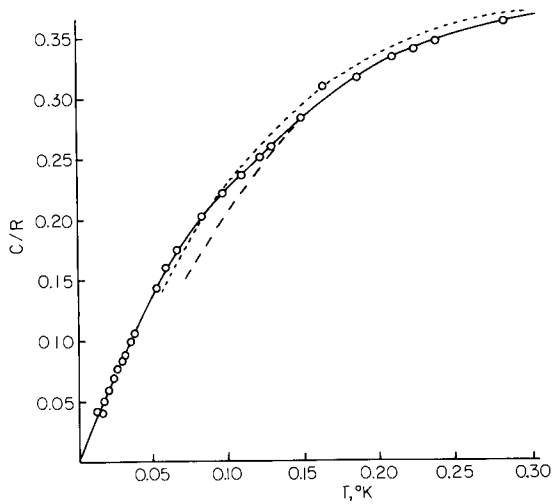


Fig. 4.13. The heat capacity, divided by the gas constant, of liquid ^{3}He. (——) Anderson *et al.* (*11*); (– –) Brewer *et al.*; (- - -) Strongin *et al.* [After Anderson *et al.* (*11*).]

Figure 4.13 illustrates the heat capacity of liquid ^{3}He, which also obeys Fermi statistics (*11*). In this (large-scale) plot, we observe that the heat capacity is roughly proportional to T below 0.04°K.

This linear dependence may be interpreted in terms of Landau's *Fermi liquid theory* (*12*). The excitations in a Fermi liquid such as liquid ^{3}He are not necessarily free from interactions, but if the interactions are switched on very slowly (adiabatically), the energy states may still be classified by the free-particle states.

If there are N_i quasi particles in state i, the entropy of the system may be given from a combinational consideration as follows:

$$S = -k \sum_i [N_i(\log N_i) + (1 - N_i) \log(1 - N_i)]. \tag{6.20}$$

We assume that the energy increase is given by

$$\delta E = \sum_i \varepsilon_i \, \delta N_i. \tag{6.21}$$

Under our assumption that the total number of quasi particles equals the number of atoms, we have

$$N = \sum_i N_i. \tag{6.22}$$

The distribution N_i that maximizes the entropy for given E and N may be determined by the method of Lagrange multipliers. The result is

$$N_i = \frac{1}{e^{(\varepsilon_i - \mu)/kT} + 1}, \tag{6.23}$$

i.e., the Fermi distribution, although ε_i may depend on temperature in the present case.

The heat capacity at very low temperatures, at which only very few quasi particles exist, may be found by differentiating the internal energy with ε_i replaced by ε_i^0, such that

$$\frac{\partial U}{\partial T} = \sum \varepsilon_i^0 \frac{\partial N_i}{\partial T}. \tag{6.24}$$

Then, the evaluation of the heat capacity follows the same procedure as for the case of free fermions, and the result will be given by Eq. (6.17). If the quasi particles have effective mass m^* and $g = 2$, one finds

$$\frac{C_V}{N} = \left(\frac{V\pi}{3N}\right)^{2/3} \frac{m^*}{\hbar^2} k^2 T. \tag{6.25}$$

Comparison with experiments yields $2.88m_3$ for m^* at zero pressure, m_3 being the mass of a ^{3}He atom:

$$m_3 = 5.007 \times 10^{-24} \text{ gm}.$$

Most of the measurements of the heat capacity of liquid ^{3}He have been performed on a magnetic temperature scale T^* using powdered cerium magnesium nitrate. Down to about 3 m°K, this temperature is the same as the absolute temperature. Below 3 m°K, differences appear, but the differences have been estimated to be within a few tenths of a millidegree.

Figure 4.14 illustrates the heat capacity (*12*) of liquid ^{3}He plotted against the magnetic temperature T^*. The ordinate is the normalized heat capacity C/nR divided by T^*, where n is the number of moles. Within the experimental error, the ratio C/nRT^* varies linearly with T^* from the lowest temperature to 125 m°K. At around 125 m°K, there appears some upward turn of the data. The data may be fitted to

$$\frac{C}{nRT^*} = A - BT^*. \tag{6.26}$$

The two constants A and B are found to be 4.63 $(°K)^{-1}$ and 22.5 $(°K)^{-2}$, respectively. Above 50 m°K, the data may also be fitted to

$$\frac{C}{nRT} = A + \Gamma T^2 \log\left(\frac{T}{\theta}\right), \tag{6.27}$$

where Γ and θ are constants. There are theories that support Eq. (6.26) or Eq. (6.27) (*13*).

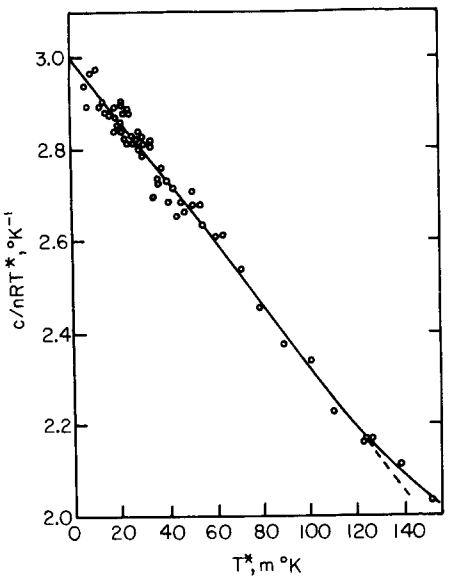


Fig. 4.14. The heat capacity of liquid ^{3}He plotted against the magnetic temperature T^*. At very low temperatures, the specific heat can be fitted to $C/nRT^* = \gamma - \beta T^*$. [After Mota *et al.* (*12*).]

4.7. Elementary Treatment of Transport Phenomena in Degenerate Gases

Not only the equilibrium properties discussed in the previous section, but also the transport properties of quantum gases can differ from those of classical gases. Using the elementary method discussed in Section 1.6, we shall study the differences (*14*).

4.7.1. The Bose Gas

Let us investigate the effect of Bose–Einstein condensation on transport phenomena. The particles in the lowest energy state have no momentum and may be considered not to transport physical quantities. The particles in the excited states have mean velocity given by

$$\bar{v} = \left(\frac{2kT}{m\pi}\right)^{1/2} \frac{\pi^2}{3G_{3/2}(1)} = 1.26\left(\frac{2kT}{m\pi}\right)^{1/2}. \tag{7.1}$$

This value is slightly smaller than that for the Maxwell gas given by Eq. (3.1) of Chapter 1. The mean free path l is inversely proportional to the density, and we may put

$$l = \frac{1}{Qn}, \tag{7.2}$$

where Q is the effective cross section and is of the order of magnitude of the square of the molecular diameter. At low temperatures, Q is about $4\pi a^2$ for rigid spherical particles and is approximately four times the classical cross section [cf. Chapter 1, Eq. (5.2)]. Thus, the viscosity is, for $T < T_0$,

$$\eta = \frac{\pi^2}{9}\left(\frac{2mkT}{\pi}\right)^{1/2} \frac{n_1}{nQG_{3/2}(1)} \tag{7.3}$$

if we use Eq. (6.3) of Chapter 1.

Similarly, the heat conductivity is calculated as follows. From Eqs. (3.19) and (4.6), the internal energy per particle is

$$\frac{U}{N_1} = \frac{3}{2}kT\frac{G_{5/2}(1)}{G_{3/2}(1)}. \tag{7.4}$$

This energy is transported by the excited particles of density n_1. Using Eq. (6.6) of Chapter 1 and Eq. (7.4), we find the following thermal conductivity:

$$\kappa = \frac{\pi^2}{3}k\left(\frac{kT}{2\pi m}\right)^{1/2} \frac{n_1}{nQ}\frac{G_{5/2}(1)}{G_{3/2}^2(1)}. \tag{7.5}$$

Since n_1/n depends on temperature, as seen from Eq. (4.5), the temperature dependences of the viscosity and of the thermal conductivity are different from those of the Maxwell gas.

4.7.2. The Fermi Gas

The average speed of fermions at 0°K is

$$\bar{v} = \frac{3}{4}\left(\frac{2\varepsilon_0}{m}\right)^{1/2}. \tag{7.6}$$

On the other hand, the mean free path of fermions at the absolute zero of temperature is infinite because all the energy levels below ε_0 are filled up and the particles cannot make transitions to different energy states. As the temperature goes up slightly above 0°K, vacant energy levels below ε_0 and excited particles appear in proportion to kT/ε_0. The mean free path is inversely proportional to the number of excited particles and also to the probability that they can make transitions, the latter being proportional to the number of vacant levels. We use

$$l = \frac{1}{nQ}\left(\frac{\varepsilon_0}{kT}\right)^2. \tag{7.7}$$

The average speed is still given by Eq. (7.6) in the first approximation. The viscosity is given by

$$\eta = \frac{1}{4}\left(\frac{\varepsilon_0}{kT}\right)^2 \frac{(2\varepsilon_0 m)^{1/2}}{Q}. \tag{7.8}$$

The viscosity is proportional to T^{-2} near 0°K.

From Eq. (6.17), the specific heat is

$$c_V = \frac{\pi^2}{2}\frac{k^2 T}{m\varepsilon_0} \quad \text{(per atom per mass)}. \tag{7.9}$$

We thus find the thermal conductivity to be proportional to T^{-1}:

$$\kappa = \frac{\pi^2}{8} k\left(\frac{\varepsilon_0}{kT}\right)\frac{(2\varepsilon_0)^{1/2}}{Qm^{1/2}}. \tag{7.10}$$

Figure 4.15 illustrates the experimental viscosity data of Becker *et al.* on

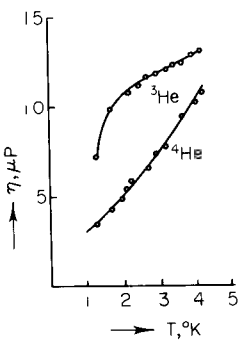


Fig. 4.15. The viscosities of ^{3}He and of ^{4}He gases. [After E. W. Becker, R. Misenta, and F. Schmeissner, *Z. Phys.* **137**, 126 (1954).]

^{3}He gas. For comparison, the viscosity of ^{4}He gas is also shown (*15*). The experiment was performed by reducing the pressure, to prevent liquefaction of the gases.

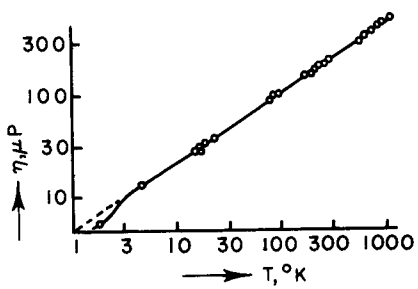


Fig. 4.16. The viscosity of ^{4}He gas over a wide temperature range. The straight line represents $\eta \quad 5.023T^{0.647}$ μP. The dashed line is the extrapolation of the straight line.

Figure 4.16 shows the viscosity of ^{4}He in a wide range of temperature (*15–17*). From the straight portion, we find

$$\eta = 5.023T^{0.647} \quad \mu\text{P}. \tag{7.11}$$

The power is larger than 0.5, indicating the role of intermolecular forces.

4.8. De Haas–van Alphen Effect

In 1930, de Haas and van Alphen found that the magnetic susceptibility of bismuth is not constant, as had been expected, but varies with the magnetic field in an oscillatory manner—*the de Haas–van Alphen effect*. This fascinating phenomenon takes place only at very low temperatures. Therefore, as Landau showed in 1930, the theoretical treatment must be based on the quantization of the free-electron states in the magnetic field. Around 1952, Lifshitz [later with Kosevich (*18*)] and, independently, Onsager (*19*) studied the magnetic havior of metals with arbitrary energy surfaces, and showed that the period of the de Haas–van Alphen oscillations is inversely proportional to the cross section of the Fermi surface in the plane normal to the field. Since then, the de Haas–van Alphen effect has been studied theoretically by many workers (*20–22*).

Due to these theories, the de Haas–van Alphen effect has become one of the most powerful tools for studying the structure of the Fermi surfaces of metals.

A free electron in a magnetic field parallel to the z direction moves freely in the z direction but makes a cyclotron motion in the xy plane, which is perpendicular to the magnetic field. Therefore, the energy is given by

$$\varepsilon_n(k_z) = \left(n + \frac{1}{2}\right)\hbar\omega_0 + \frac{\hbar^2}{2m}k_z^2, \tag{8.1}$$

where $\omega_0 = eH/mc$ is the cyclotron frequency.

The energy is quantized in the (k_x, k_y) plane. The number of states is

$$\frac{1}{4\pi^2}dk_x\,dk_y = \frac{1}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)\left\{\pi d\left[\frac{\hbar^2}{2m}(k_x^2 + k_y^2)\right]\right\}$$

$$= \frac{1}{4\pi^2}\frac{2m\pi}{\hbar^2}\hbar\omega_0 = \frac{eH}{hc} \quad \text{(per unit area)} \tag{8.2}$$

In this plane, the kinetic energy is

$$\frac{\hbar^2}{2m}(k_x^2 + k_y^2) = \left(n + \frac{1}{2}\right)\hbar\omega_0. \tag{8.3}$$

The quantity $\pi(k_x^2 + k_y^2)$ represents the cross section of the energy surface perpendicular to k_z. As we shall see shortly, the magnetic susceptibility at low temperatures is determined by $\varepsilon_0/\omega_0\hbar$. Therefore, the de Haas–van Alphen effect can be used to determine the cross section of the Fermi surface. Although we treat only the case of a spherical Fermi surface, the same conclusion holds for more general cases, making the de Haas–van Alphen effect highly useful.

The evaluation of the susceptibility will be made under the condition

$$\varepsilon_0 \gg \omega_0\hbar > kT \tag{8.4}$$

where ε_0 is the Fermi energy. This condition is necessary for the occurrence of the de Haas–van Alphen effect. If we insert numerical values, the condition on the magnetic field becomes

$$H \sim 10^4 T \quad \text{G}. \tag{8.5}$$

As we shall see, the fundamental period of oscillation ΔH is given by

$$\frac{\Delta H}{H} \sim \frac{\omega_0\hbar}{\varepsilon_0} \sim 10^{-4}. \tag{8.6}$$

Equations (8.5) and (8.6) state that the magnetic field should be very high and homogeneous and the temperature should be very low.

The Fermi energy depends on the density of the electrons and on the magnetic field. Neglecting the field dependence of the Fermi energy, we present an approximate evaluation of the susceptibility. For this purpose, we assume very low temperatures and high electron densities. The grand partition function may be approximated as follows [see Eq. (3.7)]:

$$\begin{aligned}\log \Xi &= 2\sum_i \log(1 + \lambda e^{-\beta\varepsilon_i}) \\ &\approx 2\sum_i \beta(\varepsilon_0 - \varepsilon_i), \qquad \lambda \to \infty,\end{aligned} \tag{8.7}$$

where the factor 2 represents the weight due to electronic spins. In deriving Eq. (8.7), we have used $\lambda = e^{\beta\varepsilon_0}$ and neglected 1 in the logarithm. The summation over i involves the summation over discrete Landau states n and an integration over k. Performing the integration over k_z in the interval $|k_z|^2 < [\varepsilon_0 - (n + \frac{1}{2})\hbar\omega_0]/(\hbar^2/2m)$, we find

$$\begin{aligned}\log \Xi &= \frac{2\beta VeH}{ch(2\pi)} \sum_{n=0}^{\infty} \int dk_z \left[\varepsilon_0 - \left(n + \frac{1}{2}\right)\hbar\omega_0 - \frac{\hbar^2 k_z^2}{2m}\right] \\ &= \frac{2^{5/2}}{3(2\pi^2)} (\beta\hbar\omega_0) \left(\frac{\mu_B H}{\hbar^2/2m}\right)^{3/2} V \sum_n \left[\alpha - \left(n + \frac{1}{2}\right)\right]^{3/2},\end{aligned} \tag{8.8}$$

where α is a dimensionless parameter and μ_B is the Bohr magneton:

$$\alpha = \frac{\varepsilon_0}{\hbar\omega_0}, \qquad \mu_B = \frac{e\hbar}{2mc}. \tag{8.9}$$

The summation over n can be performed by making use of

$$\begin{aligned}\sum_{n=0}^{\infty} f\left(n + \frac{1}{2}\right) &= \sum_{n=0} \int_0^{\infty} f(x)\, \delta\left(x - \left(n + \frac{1}{2}\right)\right) dx, \qquad n \geqslant 0 \\ &= \int_0^{\infty} dx\, f(x) \int_{-\infty}^{\infty} \frac{e^{2\pi i q x}}{2i \sin \pi q}\, dq \\ &= \sum_{l=-\infty}^{\infty} (-)^l \int_0^{\infty} f(x) e^{2\pi i l x}\, dx.\end{aligned} \tag{8.10}$$

In particular, for $f(x) = (\alpha - x)^{3/2}$ and for $l \neq 0$,

$$\begin{aligned}\int_0^{\alpha} (\alpha - x)^{3/2} \cos 2\pi l x\, dx &= \frac{3}{2}\frac{\alpha^{1/2}}{(2\pi l)^2} - \frac{3}{4(2\pi l)^2} \int_0^{\alpha} \frac{\cos 2\pi l x}{(\alpha - x)^{1/2}}\, dx \\ &= \frac{3}{2}\frac{\alpha^{1/2}}{(2\pi l)^2} - \frac{3\cos(2\pi l\alpha - \frac{1}{4}\pi)}{16\pi^2 2^{1/2} l^{5/2}}.\end{aligned}$$

Using these results in Eq. (8.8), we obtain

$$\begin{aligned}\frac{1}{V\beta} \log \Xi = \frac{2^{3/2}}{3\pi^2} \left(\frac{\mu_B H}{\hbar^2/2m}\right)^{3/2} \hbar\omega_0 \\ \times \left[\frac{2}{5}\alpha^{5/2} - \frac{\alpha^{1/2}}{16} - \frac{3}{2^{1/2}8\pi^2} \sum_{l=1}^{\infty} (-)^l \frac{\cos(2\pi l\alpha - \frac{1}{4}\pi)}{l^{5/2}}\right].\end{aligned} \tag{8.11}$$

The susceptibility is obtained from

$$\chi = \frac{1}{VH\beta} \left[\frac{\partial(\log \Xi)}{\partial H}\right]. \tag{8.12}$$

Several terms appear upon differentiating Eq. (8.11), but the oscillating term contributing most is given by

$$\chi_{\text{osc}} = -\frac{1}{2^{3/2}\pi^3} \left(\frac{e^2}{c\hbar}\right) \frac{\varepsilon_0}{(mc^2\mu_B H)^{1/2}} \sum_l \frac{(-)^l \sin(2\pi l\alpha - \frac{1}{4}\pi)}{l^{3/2}}. \tag{8.13}$$

Therefore, the fundamental frequency is determined by $\alpha = \varepsilon_0/\hbar\omega_0$. The term $e^2/c\hbar$ is the fine-structure constant, which is equal to 1/137.04.

In deriving Eq. (8.13), we have employed the approximation of Eq. (8.7). More generally, if such an approximation is not used and if the electrons have

spins, one obtains an expression in which a factor $\cos(\pi gl/2)/\sinh(\pi^2 kTl/\mu_B H)$ appears instead of $\mu_B H/\pi^2 kTl$:

$$\chi_{osc} = \frac{1}{2^{3/2}\pi^2}\left(\frac{e^2}{c\hbar}\right)\frac{{\varepsilon_0}^{1/2}}{(mc^2)^{1/2}}F(H), \tag{8.14}$$

$$F(H) = \frac{\pi {\varepsilon_0}^{1/2}}{(\mu_B H)^{1/2}}\frac{kT}{(\mu_B H)}\sum_l \frac{(-)^{l+1}\cos(\frac{1}{2}\pi gl)\sin(2\pi l\alpha - \frac{1}{4}\pi)}{l^{1/2}\sinh(\pi^2 kTl/\mu_B H)}.$$

Here, g is Landé's factor.

Figure 4.17 illustrates $F(H)$ in the above expression as a function of

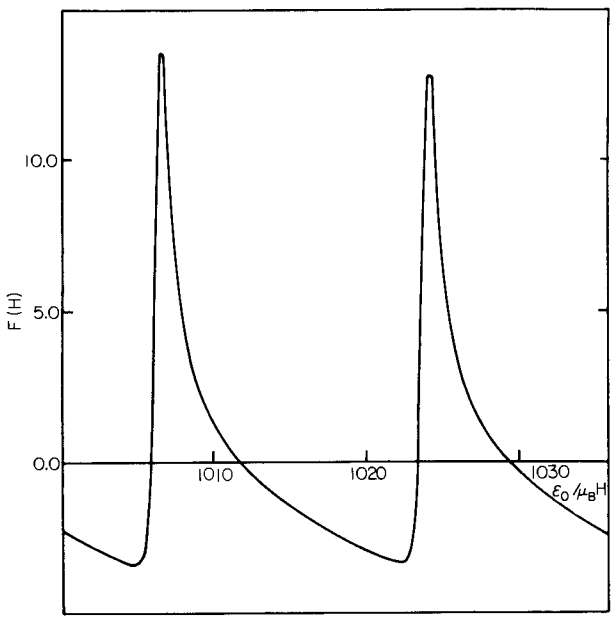


Fig. 4.17. A theoretical curve for the de Haas–van Alphen effect. The ordinate represents F in Eq. (8.14). The abscissa is $\varepsilon_0/\mu_B H$ for $T = 1.0$°K. The Fermi energy $\varepsilon_0 = {p_0}^2/2m$ corresponds to $r_s = \hbar(9\pi/4)^{1/3}/(p_0 a_0) = 4.0$, where a_0 is the Bohr radius.

$\varepsilon_0/\mu_B H$ for $g = 2$ and $T = 1$°K. The Fermi energy ε_0 is taken such that $r_s = (9\pi/4)^{1/3}\hbar/p_0 a_0$ is equal to 4.0, where the variable r_s is a dimensionless parameter which will be discussed in Chapter 14, p_0 is the Fermi momentum, and a_0 is the Bohr radius. Note in the graph that the amplitude decreases slightly with $\varepsilon_0/\mu_B H$ (*23*).

4.9.* Parastatistics

Usually, quantum statistics means either Bose or Fermi statistics. Yet another, somewhat artificial statistics, a hybrid of Bose and Fermi statistics, can be considered. In this so-called *parastatistics*, up to a finite number, say p, of particles can be in each energy state. If $p = 1$, we have the Fermi–Dirac case, and if $p = \infty$, the Bose–Einstein case.

The distribution function of particles obeying parastatistics can be obtained by the Boltzmann method discussed in Section 3.1 just as could those for the Bose and Fermi cases. Let us introduce $n_i(r)$ as the probability distribution that the ith energy state has r particles. Since each state can take up to p particles, we introduce

$$g_i = \sum_{r=0}^{p} n_i(r) \tag{9.1}$$

as a constant weight factor for the state i. The total number of particles in the state i is given by

$$N_i = \sum_{r=0}^{p} r n_i(r). \tag{9.2}$$

The most probable distribution function can be obtained by maximizing the number of ways of distributing particles in all possible energy states under the conditions that the total number of particles and the total energy are constant. Denoting the number of ways by W as in Section 3.1, we find

$$W = \prod_i \frac{g_i!}{n_i(0)!\, n_i(1)! \cdots n_i(p)!}, \tag{9.3}$$

and the restricting conditions

$$N = \sum_i N_i, \qquad U = \sum_i N_i \varepsilon_i. \tag{9.4}$$

By using the method of Lagrange multipliers and Stirling's formula, we arrive at the "equilibrium" distribution function $f_P^{(0)}(x) = n_i/g_i$ as follows:

$$f_P^{(0)}(x) = \frac{1}{e^{\alpha+x} - 1} - \frac{p+1}{e^{(p+1)(\alpha+x)} - 1}. \tag{9.5}$$

Here, $x = \varepsilon_i/kT$ and $\alpha = -\mu/kT$. The procedure of introducing the temperature and the identification of the Lagrange multiplier α are the same as before.

The para distribution function given by Eq. (9.5) has the required property:

$$f_{\mathrm{P}}^{(0)}(x) = \begin{cases} f_{\mathrm{BE}}^{(0)}(x), & p \to \infty, \\ f_{\mathrm{FD}}^{(0)}(x), & p = 1. \end{cases} \tag{9.6}$$

It also has two interesting limiting properties:

(*i*) $\alpha + x = 0$. The limit of $f^{(0)}$ is finite instead of infinite:

$$\lim_{\alpha+x\to 0} f_{\mathrm{P}}^{(0)}(x) = \tfrac{1}{2}p. \tag{9.7}$$

This is the case corresponding to the Bose–Einstein condensation.

(*ii*) $T \to 0$.

$$\lim_{T\to 0} f_{\mathrm{P}}^{(0)}(x) = \begin{cases} p, & \varepsilon < \varepsilon_0, \\ 0, & \varepsilon > \varepsilon_0, \end{cases} \tag{9.8}$$

where

$$\varepsilon_0 = \left(\frac{6n}{\pi p}\right)^{2/3} \left(\frac{h^2}{8m}\right) \tag{9.9}$$

corresponds to the Fermi energy.

The internal energy U and the number density n are given by

$$U = \frac{3}{2} NkT \frac{G_{5/2}(\lambda)}{G_{3/2}(\lambda)}. \tag{9.10}$$

$$n = \left(\frac{2\pi mkT}{h^2}\right)^{3/2} G_{3/2}(\lambda), \tag{9.11}$$

where $\lambda = e^{-\alpha}$ and

$$\begin{aligned} G_{3/2}(\lambda) &= \frac{2}{\pi^{1/2}} \int_0^\infty x^{1/2} f_{\mathrm{P}}^{(0)}(x)\, dx, \\ G_{5/2}(\lambda) &= \frac{4}{3\pi^{1/2}} \int_0^\infty x^{3/2} f_{\mathrm{P}}^{(0)}(x)\, dx. \end{aligned} \tag{9.12}$$

The relation between parastatistics and spin is analogous to the cases of ordinary statistics. There are parabosons or parafermions in accordance with integral or half-integral spins. Parabosons of order p can be antisymmetric under interchanges of up to p particles. Similarly, parafermions of order p can be symmetric under interchanges of up to p particles.

The theory of parastatistics has yet to be developed. However, it is attracting attention since there is a speculation that quarks might be parafermions of order 3 (*24*, *25*).

Problems

1. Using the probability of the occupation of any state in the grand ensemble, show that the probability distribution of an electron in an energy state at a temperature T is

$$f(E) = \frac{1}{1 + \frac{1}{2}e^{\beta(E-\mu)}}.$$

Assume that the electrostatic interaction of two electrons of opposite spins is so strong that the state is usually occupied by at most one electron.

2. Discuss the temperature dependence of the specific heat of a one-dimensional electron gas at very low temperatures.

3. Relativistic particles have the energy expression

$$\varepsilon = c[p^2 + (mc)^2]^{1/2}.$$

Using $p = mc \sinh\theta$, observe that the total number of particles, the internal energy, and the pressure of a gas of relativistic fermions (spin $\frac{1}{2}$) are given by

$$N = 8\pi\left(\frac{mc}{h}\right)^3 V \int_0^\infty \frac{\sinh^2\theta \cosh\theta\, d\theta}{\exp(-\beta\mu + \beta mc^2 \cosh\theta) + 1},$$

$$U = 8\pi\left(\frac{mc}{h}\right)^3 mc^2 V \int_0^\infty \frac{\sinh^2\theta \cosh^2\theta\, d\theta}{\exp(-\beta\mu + \beta mc^2 \cosh\theta) + 1},$$

$$p = \frac{8\pi}{3}\left(\frac{mc}{h}\right)^3 mc^2 \int_0^\infty \frac{\sinh^4\theta\, d\theta}{\exp(-\beta\mu + \beta mc^2 \cosh\theta) + 1}.$$

4. Show that the equation of state and the pressure of a completely degenerate relativistic gas are given by ($p_0 = \hbar k_0$)

$$p = \frac{\pi c}{h^3}\left[p_0(\tfrac{2}{3}p_0^2 - m^2c^2)(p_0^2 + m^2c^2)^{1/2} + (mc)^4 \sinh^{-1}\frac{p_0}{mc}\right],$$

$$U = \frac{\pi c}{h^3} V\left[p_0(2p_0^2 + m^2c^2)(p_0^2 + m^2c^2)^{1/2} - (mc)^4 \sinh^{-1}\frac{p_0}{mc}\right].$$

5. Show that the specific heat of a degenerate relativistic gas is given approximately by

$$c_V = k\frac{2\pi^{7/3}kT}{3^{1/3}hcn^{1/3}} \quad \text{per atom.}$$

6. Show that the slope of c_V of an ideal Bose gas at the condensation point is finite.

7. The viscosity of He II above 0.8°K is mainly due to phonons colliding with rotons. Assume that the phonon velocity is constant, the phonon density is proportional to T^4, and the roton density is proportional to $T^{1/2}e^{-\Delta/kT}$. Show that the phonon part of the viscosity coefficient is given by

$$\eta_p \sim \frac{T^{7/2}e^{\Delta/kT}}{\sigma(T)},$$

where $\sigma(T)$ is the cross section between phonons and rotons.

8. In the above problem, the phonon scattering may be considered to be similar to the scattering of long-wavelength sound waves from rigid spheres or the Rayleigh scattering of light by dust particles in the sky. What temperature dependence is to be expected for the cross section?
9. Find the order of the phase transition in an ideal Bose gas.
10. In a volume V and at 0°K, N free electrons interact with a uniform magnetic field H.

    (a) Determine the numbers of electrons with spins parallel and anti-parallel to the magnetic field.
    (b) Determine the magnetic susceptibility.

    The energy of a free electron in a magnetic field may be expressed by

$$\varepsilon_n = \frac{(n+\frac{1}{2})e\hbar}{mc}H + \frac{p_z^2}{2m}.$$

    For a given n and p_z, the number of energy states is given by

$$\frac{eH}{ch^2}V\,dp_z,$$

    where V is the total volume and p_z is the momentum in the direction of the magnetic field.
11. Using the above expressions, derive the partition function of a free electron in a magnetic field.
12. Show that the free energy per unit volume of a free fermion gas is

$$F_1 - n\mu = -2kT\sum_i \log(1 + e^{(\mu-\varepsilon_i)/kT}).$$

13. For a canonical partition function $Z(\beta)$, a Laplace transform and its inverse are:

$$\frac{Z(\beta)}{\beta^2} = \int_0^\infty \tilde{Z}(\varepsilon)e^{-\beta\varepsilon}\,d\varepsilon,$$

$$\tilde{Z}(\varepsilon) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} e^{\gamma\varepsilon}\frac{Z(\gamma)}{\gamma^2}\,d\gamma.$$

Show that the free energy given in the previous problem is related to $\tilde{Z}(\varepsilon)$ as follows:

$$F_1 - n\mu = \int_0^\infty \tilde{Z}(\varepsilon)\frac{\partial^2}{\partial\varepsilon^2}[-2kT\log(1 + e^{(\mu-\varepsilon)/kT})]\,d\varepsilon = 2\int_0^\infty \tilde{Z}(\varepsilon)\frac{\partial f^{(0)}}{\partial\varepsilon}\,d\varepsilon,$$

where $f^{(0)}$ is the Fermi distribution function.

REFERENCES

1. A. Einstein, *Ann. Physik.* **22**, 180, 800 (1907).
2. P. Debye, *Ann. Physik.* **39**, 789 (1912).
3. D. L. Martin, *Phys. Rev.* **141**, 576 (1966); **170**, 650 (1968).
4. F. London, *Phys. Rev.* **54**, 947 (1938).
5. J. L. Yarnell, G. P. Bendt, P. J. Bendt, and E. C. Kerr, *Phys. Rev.* **113**, 1379, 1386 (1959). For electronic heat capacity, see: J. G. Daunt, *in* "Progress in Low Temperature Physics" (C. J. Gorter, ed.). North-Holland, Amsterdam, 1957.
6. J. E. Robinson, *Phys. Rev. Lett.* **83**, 678 (1951).
7. L. Landau, *J. Phys. USSR* **5**, 71 (1941); *Soviet Physics JETP* **11**, 592 (1941).
8. L. Landau, *J. Phys. USSR* **11**, 91 (1947).
9. V. P. Peshkov, *Dokl. Akad. Nauk. SSSR* **45**, 365 (1946).
10. A. Sommerfeld, *Z. Physik.* **47**, 1 (1928).
11. A. C. Anderson, W. Reese, and J. C. Wheatley, *Phys. Rev.* **130**, 495 (1963).
12. A. C. Mota *et al.*, *Phys. Rev.* **177**, 266 (1969).
13. S. Doniach and W. Engelsberg, *Phys. Rev. Lett.* **17**, 750 (1966).
14. A. Isihara and T. Kihara, "Theory of Gases." Iwanami Publ. Co., Tokyo, 1953.
15. E. W. Becker, R. Misenta, and F. Schmeissner, *Z. Phys.* **137**, 126 (1954).
16. A. van Itterbeek, F. W. Schapink, G. J. v. d. Berg, and H. J. M. van Beeg, *Physica* **10** 465 (1943).
17. K. Fokkens *et al.*, *Physica* **30**, 2153 (1964).
18. I. M. Lifshitz and A. M. Kosevich, *J. Exp Theoret. Phys. USSR* **29**, 730 (1955).
19. L. Onsager, *Phil. Mag.* **43**, 1006 (1952).
20. R. B. Dingle, *Proc. Roy. Soc., Ser. A* **211**, 500 (1952).
21. A. B. Pippard, *Rep. 10th Solvey Congr., 1955*, p. 123.
22. J. M. Luttinger, *Phys. Rev.* **121**, 1251 (1961).
23. A. Isihara, J. Tsai, and M. Wadati, *Phys. Rev.* **3A**, 990 (1971).
24. O. W. Greeberg, *Phys. Rev. Lett.* **13**, 598 (1964).
25. H. S. Green, *Phys. Rev.* **138**, 270 (1953).

PART

II

CLASSICAL INTERACTING SYSTEMS

5

Linked Cluster Expansion

In 1927, Ursell showed in his thesis how the van der Waals equation could be derived microscopically by using a cluster expansion of the partition function of a nonideal gas. Then, around 1937, Mayer, Kahn, and Uhlenbeck and others greatly extended the theory and established the so-called linked cluster expansion theory. Similar cluster expansions are now used widely in the analysis of many-body problems in fluids and in solids.

In this chapter, we shall follow the historical development and derive first the cluster expansion and then the density expansion of the equation of state of a nonideal gas. (Those who would like to arrive at the density expansion quickly may start with Section 5.5.)

We shall then discuss the condensation phenomena and some applications of the cluster theory.

5.1. Second Virial Coefficient

The partition function of Eq. (3.8) of Chapter 3 for classical systems can be integrated over the momentum space first if the Hamiltonian is of the form of Eq. (3.1) of Chapter 3. The result is

$$Z_N = \left(\frac{2\pi mkT}{h^2}\right)^{3N/2} Q_N ;$$

$$Q_N = \frac{1}{N!}\int \cdots \int \exp\left[-\frac{\Phi(\mathbf{r}^N)}{kT}\right] d\mathbf{r}^N; \tag{1.1}$$

$$d\mathbf{r}^N = d\mathbf{r}_1\, d\mathbf{r}_2 \cdots d\mathbf{r}_N .$$

The value Q_N is determined by the molecular potentials and is most important for theoretical treatments of nonideal gases; it is called the *configurational partition function*. In the next few sections, we shall attempt to evaluate Q_N.

As an introduction to the general theory, we shall first derive the van der Waals equation of state. We adopt the pair potentials

$$\Phi(\mathbf{r}^N)=\sum_{i<j}\phi(r_{ij}), \tag{1.2}$$

and define functions f by

$$e^{-\phi(r_{ij})/kT}=1+f_{ij}. \tag{1.3}$$

Thus,

$$e^{-\Phi/kT}=\prod_{i<j}(1+f_{ij}). \tag{1.4}$$

The f-functions approach zero rather rapidly when the interparticle separations become large or when the temperature is high. Assuming that the f's are small, we expand Eq. (1.4) in powers of the f's:

$$e^{-\Phi/kT}=1+\sum_{i<j}f_{ij}+\sum_{i<j,k<l}\sum f_{ij}f_{kl}+\cdots. \tag{1.5}$$

Remembering that the f_{ij} depend only on the relative coordinates of the molecules i and j, we find

$$\int f_{ij}\,d\mathbf{r}_1\cdots d\mathbf{r}_N=V^{N-2}\int f_{ij}\,d\mathbf{r}_i\,d\mathbf{r}_j=V^{N-1}\int f(r)4\pi r^2\,dr. \tag{1.6}$$

We estimate some higher terms by products of similar integrals, and use $N^s/s!$ for ${}_NC_s$. Thus,

$$Q_N=\frac{1}{N!}V^N\left(1+\frac{N}{V}b_2\right)^N, \tag{1.7}$$

where b_2 is given by

$$b_2=\tfrac{1}{2}\int_0^\infty f(r)4\pi r^2\,dr. \tag{1.8}$$

Here, the boundary of the integral has been extended to infinity since $f(r)$ goes to zero very rapidly for large r. The term $-b_2$ is called the *second virial coefficient*.

For Q_N of Eq. (1.7), the equation of state may be approximated by

$$p=\frac{NkT}{V}\left[1-\frac{b_2}{v}\right]\cdot\left(v=\frac{V}{N}\right) \tag{1.9}$$

It is necessary to know the potential function in order to calculate p from this equation. In the following, we shall evaluate b_2 for a few simple potentials.

5.1.1. Rigid Spherical Molecule with a Small Attractive Force

Let us give the following form for the molecular potential:

$$\begin{aligned}\phi(r)&=-\phi_0\left(\frac{r_0}{r}\right)^m \quad &\text{for}\quad \infty>r\geqslant r_0\\ &=\infty \quad &\text{for}\quad 0\leqslant r<r_0.\end{aligned} \tag{1.10}$$

Then, we find approximately

$$b_2=\frac{a}{kT}-b, \tag{1.11}$$

where

$$b=\frac{2\pi}{3}r_0{}^3=4v_0,\qquad a=\frac{12}{m-3}\phi_0 v_0, \tag{1.12}$$

v_0 being the molecular volume. Introducing Eq. (1.11) into Eq. (1.9), we find that our approximate equation of state corresponds to the van der Waals equation

$$\left(p+\frac{a}{v^2}\right)(v-b)=kT \tag{1.13}$$

when the latter is expanded in powers of $1/v$ to order $1/v^2$. Equation (1.12) gives meaning to the parameters a and b in the van der Waals equation.

5.1.2. The Lennard-Jones Model

The Lennard-Jones potential is defined by

$$\phi(r)=\phi_0\left[\left(\frac{r_0}{r}\right)^{12}-2\left(\frac{r_0}{r}\right)^6\right]. \tag{1.14}$$

Here, $-\phi_0$ represents the minimum value of ϕ, which occurs at r_0.

The second virial coefficient is obtained by calculating

$$b_2=\frac{2\pi}{3}\int_0^\infty\frac{1}{kT}\frac{d\phi}{dr}e^{-\phi(r)/kT}r^3\,dr. \tag{1.15}$$

The evaluation may be made by expanding the attractive part of the potential. The result is

$$b_2=\frac{2\pi}{3}r_0{}^3I\left(\frac{\phi_0}{kT}\right), \tag{1.16}$$

where

$$I(x)=\frac{1}{4}\sum_{n=0}^{\infty}\frac{2^n}{n!}\Gamma\left(\frac{2n-1}{4}\right)x^{(2n+1)/4}. \tag{1.17}$$

The temperature at which b_2 vanishes is called the *Boyle temperature* We find that this temperature is

$$T_B=3.44\frac{\phi_0}{k}. \tag{1.18}$$

Equation (1.16) predicts that $b_2/{r_0}^3$ is a universal function of ϕ_0/kT for every gas with a Lennard-Jones potential. This is an example of the *law of corresponding states* (*1*). This prediction has been confirmed by experiments. In general, $-b_2$ becomes negative at low temperatures.

5.1.3.* The Dumbbell Molecule

As a simplified model of a diatomic molecule, we introduce a rigid dumbbell composed of two spheres of radius r_0 at a distance l apart. The interaction depends on the associated angles, and b_2 is obtained by integrating f over all relative distances and angles of two bumbbells. The second virial coefficient is obtained by counting all the configurations in which the two dumbbells are in contact with each other (then $f=-1$, otherwise $f=0$).

The rigorous second virial coefficient for rigid dumbbell molecules is illustrated in Fig. 5.1 (*2*). The ordinate is $-b_2/4v_d$, where v_d is the volume of a dumbbell molecule. The abscissa is l/r_0. It is interesting to observe that $-b_2$

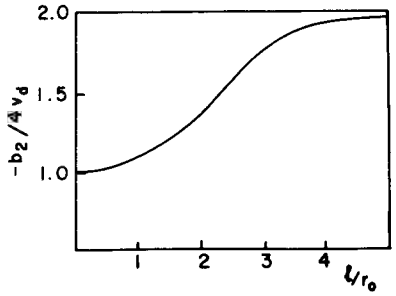


Fig. 5.1. The exact second virial coefficient for rigid dumbbell molecules.

reduces smoothly to the second virial coefficient of a hard-sphere gas in the limit of small l and to twice this coefficient in the limit of large l. In the latter limit, $-b_2/4v_d$ approaches 2, which is simply the number of spheres in a dumbbell.

5.1.4.* Rigid, Convex Molecules

The second virial coefficient of rigid, convex molecules has been evaluated quite generally and rigorously in the following form (*3*):

$$B=v_0\left(1+\frac{\rho s}{v_0}\right), \tag{1.19}$$

where v_0 is the volume of a molecule, s is its surface area, and ρ is its average radius, defined by

$$\rho=\frac{1}{4\pi}\int\frac{R_1+R_2}{2}\,d\omega, \tag{1.20}$$

where R_1 and R_2 are the principal radii of curvature and $d\omega$ is the surface element of a unit sphere.

The above formula holds even if the molecular shape is angular. The proof involves mathematics somewhat too complicated to be presented in this book, but the general method to evaluate the average radius of such molecules has been given.

For the sake of comparison with the case of spherical molecules, it is convenient to rewrite Eq. (1.19) as follows:

$$B=4v_0f; \qquad f=\frac{1}{4}\left(1+\frac{\rho s}{v_0}\right). \tag{1.21}$$

The coefficient f is unity when the molecules are rigid spheres. Generally, f is larger than one for nonspherical molecules. In the case of a cylinder or an ellipsoid of revolution, it increases with the cylindrical length or the major ellipsoidal axis when these quantities are very large.

In fact, the second virial coefficient for rigid ellipsoidal molecules of rotational symmetry is given rigorously by

$$f=\frac{1}{4}\left[1+\frac{3}{4}\left(1+\frac{1}{(1-\varepsilon^2)^{1/2}}\frac{\sin^{-1}\varepsilon}{\varepsilon}\right)\left(1+\frac{1-\varepsilon^2}{2\varepsilon}\log\frac{1+\varepsilon}{1-\varepsilon}\right)\right], \tag{1.22}$$

where ε is the eccentricity.

Table 5.1 shows the coefficient f and other relevant quantities for rigid molecules.

TABLE 5.1

Quantities Related to the Second Virial Coefficients for Rigid Molecules

	Sphere (radius 1)	Cube (side 1)	Regular tetrahedron (side 1)	Regular octahedron (side 1)	Circular cylinder (length l, radius a)	Circular disk (radius 1)	Line segment (length 1)
v_0	$4\pi/3$	1	$\sqrt{2}/12$	$\sqrt{2}/3$	$\pi a^2 l$	0	0
s	4π	6	$\sqrt{3}$	$2\sqrt{3}$	$2\pi a(a+l)$	2π	0
ρ	1	3/4	$(3/2\pi)\tan^{-1}\sqrt{2}$	$(3/\pi)\cot^{-1}\sqrt{2}$	$(\pi a + l)/4$	$\pi/4$	1/4
f	1	1.375	1.926	1.330	$\dfrac{\pi a^2 + (\pi+3)al + l^2}{8al}$	—	—

5.1.5.* Quantum Gases

The second virial coefficient of a quantum gas is given by

$$B = -b_2 . \tag{1.23}$$

We use λ for the thermal de Broglie wavelength $(h^2/2\pi mkT)^{1/2}$ and b_2, the second cluster integral, is given by

$$b_2 = \frac{1}{2V}\iint d\mathbf{r}_1\, d\mathbf{r}_2(\langle 1, 2 | e^{-\beta H_2} | 1, 2\rangle - \langle 1| e^{-\beta H_1}|1\rangle\langle 2 | e^{-\beta H_1}|2\rangle), \tag{1.24}$$

where H_1 and H_2 are, respectively, the Hamiltonians of one particle and of a pair of particles. The angular bracket means a matrix element taken with properly symmetrized wave functions.

In the absence of interactions, the cluster coefficient b_2 is given by

$$b_2{}^0 \begin{cases} =0 & \text{(classical)}, \\ =2^{-5/2}\lambda^3 & \text{(bosons)}, \\ =-2^{-5/2}\lambda^3 & \text{(fermions)}. \end{cases} \tag{1.25}$$

The existence of the second virial coefficient for an ideal quantum gas corresponds to the equation of state given by Eq. (3.17) of Chapter 4.

In the presence of interactions, the matrix elements in Eq. (1.24) may be given by separating the coordinates of a pair of particles into center-of-mass and relative coordinates. The former give rise to a factor V, and the latter lead to a factor determined by the radial wave functions. We obtain

$$b_2 - b_2{}^0 = 2^{3/2}\lambda^3\left\{\sum_l (2l+1)(\exp\beta\varepsilon_l) + \frac{1}{\pi}\sum_l (2l+1) \times \int_0^\infty dk\, \frac{\partial\delta_l(k)}{\partial k}\exp -\frac{\beta\hbar^2k^2}{m}\right\}, \tag{1.26}$$

where the $\delta_l(k)$ represent phase shifts. For classical particles, the summation over l extends over all positive integers and zero. It is over all even integers and zero for bosons, and over all positive odd integers for fermions. The result (1.26) was obtained first by Beth and Uhlenbeck (*4*).

5.1.6. The Joule–Thomson Effect

The second virial coefficient of a gas can be determined by measuring its pressure. An alternative method uses the so-called *Joule–Thomson effect*: the temperature change in an adiabatic expansion of the gas.

The temperature change of a gas due to an adiabatic expansion is given by $(\partial T/\partial p)_H$, where H is the enthalpy, defined by

$$H = U + pV. \tag{1.27}$$

Taking the total derivatives of both sides of this equation and making use of Eq. (2.6) of Chapter 3, we find

$$dH = T\,dS + V\,dp. \tag{1.28}$$

Thus,

$$\left(\frac{\partial T}{\partial p}\right)_H = -\frac{(\partial H/\partial p)_T}{(\partial H/\partial T)_p} = \frac{T(\partial V/\partial T)_p - V}{C_p}, \tag{1.29}$$

where C_p is the heat capacity at constant pressure.

Formula (1.29) is general. In particular, if the equation of state is of the form

$$V = \frac{NkT}{p} + B(T),$$

the formula gives

$$\left(\frac{\partial T}{\partial p}\right)_H = \frac{T^2}{C_p}\frac{d(B/T)}{dT}. \tag{1.30}$$

This formula can be used to determine the second virial coefficient.

5.2. Cluster Expansion

5.2.1. The Canonical Ensemble Method

Let us consider deriving the equation of state more generally than in the previous section. In view of Eqs. (1.1) and (1.5), we define the so-called *cluster integrals* as follows (5):

$$\begin{aligned} b_1 &= \frac{1}{V}\int d\mathbf{r}_1 = 1, \\ b_2 &= \frac{1}{2!V}\iint f_{12}\,d\mathbf{r}_1\,d\mathbf{r}_2, \\ b_3 &= \frac{1}{3!V}\iiint (f_{31}f_{21} + f_{32}f_{31} + f_{32}f_{21} + f_{21}f_{32}f_{31})\,d\mathbf{r}^3, \\ &\vdots \\ b_l &= \frac{1}{l!V}\iint\cdots\int \sum\prod f_{ij}\,d\mathbf{r}^l. \end{aligned} \tag{2.1}$$

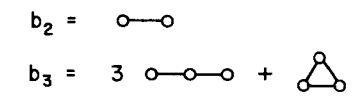


Fig. 5.2. Graphs illustrating the cluster integrals b_2 and b_3.

The last integrand consists of a sum of products of f's of particles which are connected with each other by *f-bonds*. The integrands of $b_2, b_3, \ldots$ may be illustrated by "connected graphs" of 2, 3, ... molecules as in Fig. 5.2, where the dots and the bonds represent, respectively, the molecules and the f-bonds. The coefficient b_l is obtained by considering all the diagrams of l molecules in which each molecule is connected to the other molecules by at least one bond.

In terms of these b_l, the configurational partition function is expressed as follows:

$$Q_N = \sum_{m_l}\prod_l \frac{(Vb_l)^{m_l}}{m_l!}, \tag{2.2}$$

where the summation and the product go over integers l and m_l that satisfy the relation

$$\sum_l l m_l = N. \tag{2.3}$$

Equation (2.2) is derived by observing that the total number of ways of distributing N unlabeled molecules into m_l identical clusters of l molecules is given by

$$\frac{N!}{\prod m_l!(l!)^{m_l}}$$

and that each of the distributions contributes to Q_N by the amount

$$\frac{1}{N!}\prod(l!Vb_l)^{m_l}.$$

Multiplication of the above two factors gives Eq. (2.2).

The partition function given by Eq. (2.2) with the subsidiary condition of Eq. (2.3) may be expressed in terms of a complex integral as follows:

$$\begin{aligned} Q_N &= \frac{1}{2\pi i}\oint \sum_{m_l}\prod_l \frac{(Nvb_l z^l)^{m_l}}{m_l!}\frac{dz}{z^{N+1}} \\ &= \frac{1}{2\pi i}\oint \frac{e^{NvG_0(z,\,T;\,N)}}{z^{N+1}}\,dz \end{aligned} \tag{2.4}$$

where

$$G_s(z, T; N) = \sum_{l=1}^{N} l^s b_l z^l. \tag{2.5}$$

For large N, the integral may be evaluated by the method of steepest descent. The saddle point is determined from the condition that the derivative of the coefficient of N in the exponent in Eq. (2.4) vanishes:

$$v\frac{dG_0}{dz}-\frac{1}{z}=0. \tag{2.6}$$

That is,

$$G_1(z, T; N)=\frac{1}{v}=n, \tag{2.7}$$

where n is the density.

Equation (2.6) can have several roots z_i, $i = 1, 2, \ldots$. At the saddle points, the second derivative of the exponent of Eq. (2.4) must be positive. The derivative is given by

$$\begin{aligned} v\frac{d^2G_0}{dz^2}+\frac{1}{z^2} &= v\sum l^2 b_l z^{l-2} \\ &= vz^{-2}G_2(z, T; N), \end{aligned} \tag{2.8}$$

Assuming that there exists just one saddle point that gives the maximum contribution to Q_N, we arrive at

$$\begin{aligned} Q_N &= \frac{\exp NvG_0}{2\pi i z^{N+1}}\oint \exp\left[\frac{NvG_2(\xi-z)^2}{2z^2}\right]d\xi \\ &= \frac{\exp NvG_0}{z^N(2\pi NvG_2)^{1/2}}, \end{aligned} \tag{2.9}$$

where G_0 and G_2 are evaluated at the saddle point.

Neglecting terms of order $1/N$, the equation of state is obtained as follows:

$$\begin{aligned} p &= \frac{kT}{N}\frac{\partial(\log Q_N)}{\partial v} \\ &= kT\frac{\partial}{\partial v}(vG_0-\log z) \\ &= kTG_0 + kTv\sum_{l=1}^{N}\frac{\partial b_l}{\partial v}z^l \\ &= kT\sum_{l=1}^{N}(vb_l)'z^l, \end{aligned} \tag{2.10}$$

where the prime indicates a derivative. Equation (2.10) is called the *cluster expansion* of the equation of state. In this expansion, the pressure is determined by a series of connected diagrams. The right-hand terms of Eq. (2.10) involve the terms coming from the volume dependence of b_l. However, this dependence is small except for the condensation range; and we shall approximate Eq. (2.10) as

$$\begin{aligned} p &= kTG_0 \\ &= kT\sum_{l=1}^{N} b_l z^l. \end{aligned} \tag{2.11}$$

Upon differentiation, we obtain

$$\frac{\partial p}{\partial v}=-\frac{kT}{v^3}\frac{1}{G_2(z, T; N)}. \tag{2.12}$$

Since at the saddle point $G_2 > 0$, we have $\partial p/\partial v < 0$, as it should be.

Problem

Show that, if b_l depends on v, Eq. (2.12) should be replaced by

$$\begin{aligned} \frac{1}{kT}\frac{\partial p}{\partial v} &= -\frac{1}{v^3G_2}\left[1+\sum_{l}^{N} l(v^2b_l')z^l\right]^2 \\ &\quad +\frac{1}{v}\sum_{l=1}^{N} z^l(v^2b_l')'. \end{aligned} \tag{2.13}$$

5.2.2. The Grand Canonical Ensemble Method

When the volume dependence of the cluster integrals is negligible and the series defined by

$$G(z, T; \infty)=\sum_{l=1}^{\infty} b_l z^l \tag{2.14}$$

is convergent, we obtain the grand partition function as follows:

$$\begin{aligned} \Xi &= \Xi(T, V; z) \\ &= \sum_{M=0}^{\infty} Q_M z^M \\ &= \exp\left(V\sum_{l=1}^{\infty} b_l z^l\right). \end{aligned} \tag{2.15}$$

Thus, use of the formulas (5.6) and (5.14) of Chapter 3 gives

$$N = \left[\frac{\partial(\log \Xi)}{\partial(\log z)}\right]_{T,V} \rightarrow N = V\sum_{l=1}^{\infty} l b_l z^l, \tag{2.16}$$

$$pV = kT \log \Xi \qquad \rightarrow p = kT\sum_{l=1}^{\infty} b_l z^l. \tag{2.17}$$

Comparing Eqs. (2.16) and (2.17) with the first and the first of Eqs. (4.11) of Chapter 3, we find that z plays the role of λ. That is, z is the activity (page 61). Upon eliminating z from these two expressions, we can determine the pressure as a function of volume. Equations (2.16) and (2.17), derived from the grand partition function, correspond to Eqs. (2.7) and (2.11), obtained from the canonical partition function.

It should be remarked, however, that the above grand ensemble treatment is based on the assumptions that the cluster integrals b_l are independent of volume and that the infinite cluster series defined by Eq. (2.14) is convergent. Also, regardless of the convergent character of the cluster series, we may not be able to put an infinite number of particles in a finite vessel. For instance, if the molecules are rigid spheres, there must be a maximum number $N_{\max}$ of molecules possible in a vessel. The appropriate grand partition function in such a case is

$$\Xi = \sum_{M=0}^{N_{\max}} Q_M z^M. \tag{2.18}$$

On the other hand, $\log \Xi$ is an infinite series in z and may be expressed in terms of new coefficients b_l^+ which are generally different from b_l except for the case $N_{\max} = \infty$:

$$\frac{1}{V}\log \Xi = \sum_{l=1}^{\infty} b_l^+ z^l. \tag{2.19}$$

For $l < N_{\max}$, the b_l^+ are given by the Thiele semiinvariants as follows:

$$\begin{aligned} Vb_1^+ &= Q_1, \\ Vb_2^+ &= Q_2 - \tfrac{1}{2}Q_1^2, \\ Vb_3^+ &= Q_3 - Q_1Q_2 + \tfrac{1}{3}Q_1^3, \\ &\cdots. \end{aligned} \tag{2.20}$$

In general,

$$Vb_l^+ = \sum (-)^{\sum_l n_l - 1}\left(\sum_l n_l - 1\right)! \prod \frac{Q_l^{n_l}}{n_l!}, \tag{2.21}$$

where the first sum and the products correspond to all sets of positive integers or zero satisfying

$$\sum_i i n_i = l.$$

Conversely, we obtain from Eq. (2.20),

$$Q_N = \sum_{m_l} \prod_l \frac{(Vb_l^+)^{m_l}}{m_l!}, \tag{2.22}$$

which is simply Eq. (2.2). Thus, for $l < N_{\max}$, $b_l^+ = b_l$. However, for $l > N_{\max}$, the b_l's are not given by the same invariants [simply because the series (2.18) is finite] and there is no Q_M for $M > N_{\max}$. In other words, $\lim_{l\to\infty} b_l^+$ may be very different from the b_l for large l.

5.2.3. Conventional Interpretation

A relatively simple interpretation may be given to the cluster series for pressure and density by considering molecular configurations as a function of density and temperature. At high temperatures and low densities, gases are more or less ideal. As the temperature is reduced and as the density is increased, two-molecule clusters will be formed, followed by three-, four-molecule clusters, etc. Finally, the liquid state will be reached when all the molecules are bound together in a macroscopic cluster.

In equilibrium fluids, a statistical distribution of clusters of molecules may be introduced as a function of temperature and density. The molecules in a connected diagram are interacting and form what may be called a "pseudomolecule." Of course, none of the pseudomolecules thus defined are rigid. Yet, statistically, there may be a certain distribution of clusters or pseudomolecules at a given temperature and density, and the gas may be considered to be equivalent to a mixture of these pseudomolecules. In our conventional interpretation, this mixture is an ideal gas mixture because there are no interactions between the pseudomolecules.

Let us assume that we have m_l clusters of l molecules. Clearly, $\sum lm_l = N$ as in Eq. (2.3). According to Henry's law for an ideal gas mixture, we have the equation of state

$$pV = kT\sum m_l \tag{2.23}$$

for the system. The number m_l is proportional to the probability of finding cluster molecules of size l. In the grand ensemble of a gas, the relative probability of an l-particle system is given by $z^l Q_l$. The configurational partition function Q_l does not correspond to a particular configuration of the molecules,

namely, a particular connected graph. On the other hand, m_l corresponds to the configurations where the l particles are connected. Therefore, m_l may be expressed, in a form similar to $z^l Q_l$, as follows:

$$m_l = V b_l z^l, \tag{2.24}$$

where the volume V is introduced only for mathematical convenience. Note that V is a constant independent of l and that Vb_l plays the role of the probability of finding an l-particle cluster. Equations (2.23) and (2.24) are identical to Eqs. (2.17) and (2.16). Also, Eq. (2.24) is consistent with Eq. (2.3). From Eq. (2.23), we may accept the interpretation that the clusters behave as if they are "molecules."

Nevertheless, this interpretation is only a convention, since the b's are not necessarily positive.

Problem

Show that m_l of Eq. (2.24) can be obtained by maximizing the terms in Q_N of Eq. (2.2) with respect to m_l with Eq. (2.3) as a constraint.

5.3. Virial Expansion

The cluster expansion of the equation of state is expressed in terms of the activity variable z, which is determined from the condition that the total number of molecules is N. By solving Eq. (2.16) by a successive approximation, we can determine z as a function of v and, from Eq. (2.17), the pressure as a function of volume, but the following method is more sophisticated. The volume dependence of the cluster integrals will be neglected in the following treatment.

We notice in Eq. (2.7) that, for large v, z should be small. The first approximation is clearly $z = 1/v$, since $b_1 = 1$. Thus, let us assume a form for z:

$$\begin{aligned} z &= w \exp\left(-\sum_{k=1}^{\infty} \beta_k w^k\right), \\ w &= \frac{1}{v} = n. \end{aligned} \tag{3.1}$$

We shall explain later the meaning of the coefficients β_k in this expression for z. At present, we assume Eqs. (3.1) and proceed to find the consequences of our assumption.

Since $w = n = \sum l b_l z^l$ and since $l^2 b_l$ is the coefficient of z^l in the expansion of $z\, dw/dz$, we have

$$l^2 b_l = \frac{1}{2\pi i} \oint \frac{dz}{z^l} \frac{dw}{dz} = \frac{1}{2\pi i} \oint \frac{\exp(l \sum_k \beta_k w^k)}{w^l}\, dw. \tag{3.2}$$

From the last expression, we find that $l^2 b_l$ is the coefficient of w^{l-1} of the expansion of $\exp(l \sum_k \beta_k w^k)$ in powers of w. Thus,

$$l^2 b_l = \sum_{m_k} \prod_k \frac{(l\beta_k)^{m_k}}{m_k!}, \tag{3.3}$$

where the summation and product are subject to

$$\sum k m_k = l - 1. \tag{3.4}$$

Equation (3.3) gives the b_l as functions of the β_k. We notice the similarity between the two expansions: Q_N expanded in b_l and b_l in β_k. The b_l represent connected diagrams. As will be shown later, the β_k correspond to *irreducible diagrams*, and are called *irreducible integrals*. From Eq. (3.3), the first few β_l are obtained from

$$\begin{aligned} b_1 = 1; \quad b_2 = \tfrac{1}{2}\beta_1; \quad b_3 = \tfrac{1}{2}\beta_1^2 + \tfrac{1}{3}\beta_2, \\ b_4 = \tfrac{2}{3}\beta_1^3 + \beta_1\beta_2 + \tfrac{1}{4}\beta_3, \quad \ldots. \end{aligned} \tag{3.5}$$

A close examination of Eq. (3.5) will show that the β_k indeed correspond to integrals that cannot be reduced to simpler integrals.

We can now express p as a function of v as follows:

$$\begin{aligned} \sum_{l=1}^{\infty} b_l z^l = \int_0^z \frac{dz}{z} \sum_1^{\infty} l b_l z^l = \int_0^w \frac{dw}{w} (1 - \sum k\beta_k w^k) w \\ = w - \sum \frac{k}{k+1} \beta_k w^{k+1}. \end{aligned} \tag{3.6}$$

Since $w = 1/v$, we obtain

$$\begin{aligned} \frac{pv}{kT} &= 1 - \sum \frac{k}{k+1} \frac{\beta_k}{v^k} \\ &= 1 + \frac{B}{v} + \frac{C}{v^2} + \cdots. \qquad (v = V/N) \end{aligned} \tag{3.7}$$

$B = -\beta_1/2$, $C = -2\beta_2/3$, ... are called, respectively, the second, the third, ... *virial coefficients*.

Problems

1. Show that

(*i*) the heat capacity at constant volume is given by

$$C_V = Nk\left[\frac{3}{2} + \sum \frac{1}{k+1}\left(\frac{\partial}{\partial T} T^2 \frac{\partial \beta_k}{\partial T}\right) v^{-k}\right]; \tag{3.8}$$

(*ii*) the third virial coefficient is given by

$$C = -\tfrac{1}{3}\int\int f(r_{12}) f(r_{13}) f(r_{23})\, d\mathbf{r}_1\, d\mathbf{r}_2 . \tag{3.9}$$

2. Using the Fourier transform

$$f_{ij}(r_{ij}) = (2\pi)^{-3}\int \tilde{f}(k) \exp i\mathbf{k}\cdot\mathbf{r}_{ij}\, d\mathbf{k}, \tag{3.10}$$

express the integrals of Eq. (3.9) in **k**-space.

5.4. Irreducible Integrals

The cluster integrals correspond to connected graphs. Some of the terms in the cluster integrals can be reduced to simpler integrals. For instance, let us consider the integral

$$J = \frac{1}{V}\int f_{12} f_{23}\, d\mathbf{r}_1\, d\mathbf{r}_2\, d\mathbf{r}_3 . \tag{4.1}$$

We note that the only function that depends on $\mathbf{r}_3$ is f_{23}, which may be assumed to be a rapidly decreasing function of r_{23}. Let us use $d\mathbf{r}_{23}$ instead of $d\mathbf{r}_3$ and integrate f_{23} first. This integration may be carried out in infinite space for f_{23} corresponding to short-range forces. The result is independent of the position $\mathbf{r}_2$. The integration of the function f_{12} over $\mathbf{r}_2$ gives the same value independent of the position $\mathbf{r}_1$. Finally, the integration over $\mathbf{r}_1$ gives the volume V. Thus, the entire integral is reduced to

$$J = \frac{1}{V}\int f_{12} f_{23}\, d\mathbf{r}_1\, d\mathbf{r}_2\, d\mathbf{r}_3 = \left[\int f(r)\, d\mathbf{r}\right]^2$$
$$= \beta_1^{\,2}. \tag{4.2}$$

By similar considerations, the cluster integrals are decomposed as in Eq. (3.5). The β_s are irreducible integrals in the sense that further reductions of the integrals are not possible. Explicitly, they are

$$\begin{aligned}
\beta_1 &= \frac{1}{1!V}\int f_{12}\, d\mathbf{r}_1\, d\mathbf{r}_2 = \int f(r)\, d\mathbf{r},\\
\beta_2 &= \frac{1}{2!V}\int f_{12} f_{23} f_{13}\, d\mathbf{r}_1\, d\mathbf{r}_2\, d\mathbf{r}_3,\\
\beta_3 &= \frac{1}{3!V}\int (3 f_{12} f_{14} f_{23} f_{34} + 6 f_{12} f_{13} f_{14} f_{23} f_{34}\\
&\quad + f_{12} f_{13} f_{14} f_{23} f_{24} f_{34})\, d\mathbf{r}_1\, d\mathbf{r}_2\, d\mathbf{r}_3\, d\mathbf{r}_4,\\
&\vdots
\end{aligned} \tag{4.3}$$

As in the case of cluster integrals, the irreducible integrals may be represented by graphs. Figure 5.3 illustrates β_3. The numbers next to the diagrams

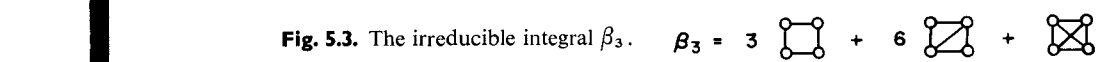


Fig. 5.3. The irreducible integral β_3.

represent the numbers of identical graphs obtained by permutation, i.e., these are the numbers appearing in β_3. As we see, all the molecules are connected by more than one bond. This type of connection is characteristic of irreducible diagrams.

The cluster integrals b_l may be expressed in terms of products of powers of irreducible integrals. We can adopt the form (6)

$$Vl!\, b_l = \sum K(n_1, n_2, \ldots) \prod_k (k!\beta_k)^{n_k} V, \tag{4.4}$$

where the summation and the products are subject to the condition $\sum k n_k = l - 1$. The unknown function $K(n_1, n_2, \ldots)$ is found if we consider $k+1$ cluster diagrams that are composed of $l_1, l_2, \ldots, l_{k+1}$ molecules. Let us choose one molecule from each of these clusters and connect them to form one more irreducible integral β_k (see Fig. 5.4). The total number of molecules is

$$l = \sum_{m=1}^{k+1} l_m ,$$

and we have

$$\frac{l!}{\prod_{m=1}^{k+1} l_m!}$$

ways of distributing the molecules among such $k+1$ clusters. Since we can

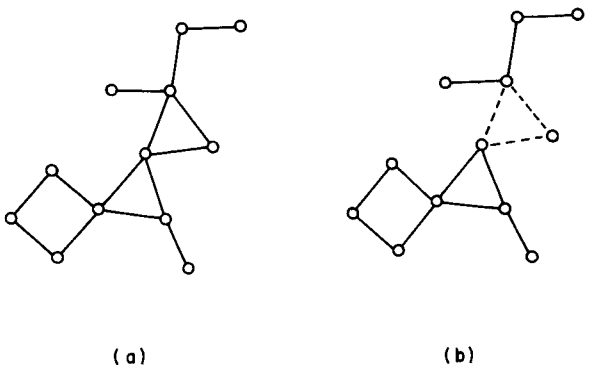


Fig. 5.4. Decomposition of a b_l.

choose any of the molecules from each of the clusters, the total number of irreducible integrals in the combined cluster must be proportional to

$$l! \prod_{m=1}^{k+1} \frac{l_m K(\ldots, n_{hm}, \ldots)}{l_m!},$$

where n_{hm} represents the number of β_h's in the mth cluster. Clearly, $\sum_m n_{hm} = n_h$ is the total number of irreducible integrals in the combined cluster. The number of β_k's in this combined cluster is $n_k + 1$. Since interchanges of molecules in the combined system should not yield different configurations, we obtain

$$(n_k + 1)(k+1)!\, K(n_1, \ldots, n_{k+1}, \ldots) = \sum l! \prod_{m=1}^{k+1} \frac{l_m K(\ldots, n_{hm}, \ldots)}{l_m!}. \quad (4.5)$$

This is a recurrence formula to determine K. However, we need not try to solve this equation. Instead, we use the generating function defined by

$$w = \sum_{l, n_k} \frac{K(n_1, n_2, \ldots)}{(l-1)!} z^l \prod_k (k!\,\beta_k)^{n_k}, \qquad \sum k n_k = l - 1, \quad (4.6)$$

and shall show that the relation between w and z determined from these equations coincides with that given by Eq. (3.1). Then, Eq. (4.4) should be identical to Eq. (3.3) and K will be automatically determined. We therefore multiply each side of

$$z^l \prod_k (k!\,\beta_k)^{n_k} = \prod_m z^{l_m} \prod_k (k!\,\beta_k)^{n_{km}}$$

by each side of Eq. (4.5). Then, we obtain

$$\frac{\partial w}{\partial \beta_k} = w^k \left(z \frac{\partial w}{\partial z} \right). \quad (4.7)$$

On the other hand, we have the relation

$$w(z, \beta_k) = \lambda w\left(\frac{z}{\lambda}; \lambda^k \beta_k \right)$$

for an arbitrary λ. Differentiating this with respect to λ and putting $\lambda = 1$, we find

$$z \frac{dw}{dz} = w + \sum_k k \beta_k \frac{\partial w}{\partial \beta_k}. \quad (4.8)$$

Using Eqs. (4.7) and (4.8), the reader will be able to show that Eq. (4.6) reduces to Eq. (3.1). Thus, the β_k introduced in Eq. (3.1) are actually the irreducible integrals defined in Eq. (4.3).

As we have seen, the irreducible integrals play an important role in converting the cluster series for the equation of state into the virial series. Therefore, we should pay attention to the fact that the finiteness of the volume has not been taken into consideration. Its effect becomes important when the total volume is reduced. For a relatively large but finite volume V, we estimate J of Eq. (4.2) as follows:

$$J \approx \beta_1^2 \left(1 + \alpha \frac{v_0}{V} \right). \quad (4.9)$$

Here, v_0 is the molecular volume and α is a factor of order unity. The factor in parentheses represents a correction due to finite volume.

5.5. Cumulant Expansion

Let us consider a function Z that can be expanded in a power series in x:

$$Z = \sum_{n=0}^{\infty} \frac{\lambda_n}{n!} x^n, \qquad \lambda_0 = 1, \quad (5.1)$$

and whose logarithm can also be expanded in powers of x:

$$\log Z = \sum_{n=1}^{\infty} \frac{\mu_n}{n!} x^n. \quad (5.2)$$

The coefficients μ_n are known as the Thiele *semiinvariants* or cumulants and are given by

$$\begin{aligned}\mu_1 &= \lambda_1,\\ \mu_2 &= \lambda_2 - \lambda_1^2,\\ \mu_3 &= \lambda_3 - 3\lambda_1\lambda_2 + 2\lambda_1^3,\\ \mu_4 &= \lambda_4 - 4\lambda_1\lambda_3 - 3\lambda_2^2 + 12\lambda_1^2\lambda_2 - 6\lambda_1^4,\\ &\vdots\end{aligned} \tag{5.3}$$

In general,

$$\mu_j = \sum_{n_i} (-)^{\sum n_i - 1}(\textstyle\sum n_i - 1)! \prod_i \left[\frac{(\lambda_{n_i}/i!)^{n_i}}{n_i!}\right], \tag{5.4}$$

where the summation of products is over all sets of integers that satisfy $\sum_i i n_i = j$.

In particular, if

$$Z = \int e^{-\beta\phi(r)} f(r)\, d\mathbf{r} = \langle e^{-\beta\phi(r)}\rangle \tag{5.5}$$

and if

$$\int f(r)\, d\mathbf{r} = 1, \tag{5.6}$$

we find by expanding $e^{-\beta\phi}$ and integrating term by term that

$$\lambda_n = \langle(-\phi)^n\rangle. \tag{5.7}$$

μ_n is the nth moment, which is free from products of uncorrelated lower order moments. Thus, we introduce the notation

$$\mu_n = \langle\lambda_n\rangle_c, \tag{5.8}$$

which means that μ_n represents the correlated part of the moment λ_n. We can then write

$$\log Z = \langle e^{-\beta\phi} - 1\rangle_c. \tag{5.9}$$

The right-hand side can be computed by expanding $e^{-\beta\phi}$ and operating with $\langle\cdots\rangle_c$ to each term.

The above consideration for $Z(x)$ can easily be generalized to a function of many variables. Let us consider the configurational partition function

$$Q = \frac{1}{N!}\int e^{-\beta\Phi}\, d\mathbf{r}^N. \tag{5.10}$$

If we introduce an averaging notation

$$\langle\cdots\rangle = \frac{1}{V^N}\int\cdots d\mathbf{r}^N, \tag{5.11}$$

Q may be expressed by

$$Q = \frac{V^N}{N!}\langle e^{-\beta\Phi}\rangle. \tag{5.12}$$

Thus, using Stirling's formula, we obtain

$$\begin{aligned}\log Q &= N[(\log v) + 1 + W],\\ NW &= \langle e^{-\beta\Phi} - 1\rangle_c.\end{aligned} \tag{5.13}$$

In particular, if $\Phi = \sum \phi_{ij}$, the second-order cumulant is of the form

$$\langle\Phi^2\rangle - \langle\Phi\rangle^2 = \sum\sum(\langle\phi_{ij}\phi_{kl}\rangle - \langle\phi_{ij}\rangle\langle\phi_{kl}\rangle).$$

Only when $(i,j) = (k,l)$ does the contribution on the right-hand side not vanish. We note that

(*i*) $\langle\phi_{ij}\phi_{kl}\rangle = \langle\phi_{ij}\rangle\langle\phi_{kl}\rangle$,
(*ii*) $\langle\phi_{ij}\phi_{jl}\rangle = \langle\phi_{ij}\rangle\langle\phi_{jl}\rangle$.

The first case corresponds to a nonconnected pair of graphs. The second case represents a connected graph, but it is reducible. In both cases, the cumulants vanish. Generalizing this observation, we arrive at an important cumulant property:

CUMULANT THEOREM. *Only irreducible diagrams contribute to* $\langle e^{-\beta\Phi} - 1\rangle_c$.

Since the function W is obtained by volume integrals, it is a function of the density $n = N/V$. Therefore,

$$\frac{p}{kT} = n - n^2\frac{\partial W}{\partial n}. \tag{5.14}$$

Let us consider constructing W by expanding the exponential function. For $\Phi = \sum_{i<j}\phi_{ij}$, the first term is just $\langle -\beta\phi_{12}\rangle$ multiplied by the number of pairs. The second term comes from three particles. The only contribution is from the terms of the form $\langle\phi_{12}\phi_{23}\phi_{31}\rangle$, and the total number is ${}_NC_3$. Thus,

$$\begin{aligned}{}_NC_2\langle\phi_{12}\rangle &= \frac{Nn}{2}\int\phi_{12}\, d\mathbf{r},\\ {}_NC_3\langle\phi_{12}\phi_{23}\phi_{31}\rangle &= \frac{Nn^2}{3!}\int\phi_{12}\phi_{23}\phi_{31}\, d\mathbf{r}_2\, d\mathbf{r}_3,\end{aligned} \tag{5.15}$$

where n is the number density. In general, we find

$$W = \sum \frac{n^k}{(k+1)!} \sum \int_{k+1} \prod_{\geq i > j \geq 1} \frac{(-\beta\phi_{ij})^{\nu_{ij}}}{\nu_{ij}!} d\mathbf{r}_2 \cdots d\mathbf{r}_{k+1}, \tag{5.16}$$

where the second summation is over all graphs of $(k+1)$ vertices (particles). Equation (5.16) can be expressed in terms of Mayer's f-functions:

$$\sum_{\nu=1}^{\infty} \frac{(-\beta\phi_{12})^{\nu}}{\nu!} = e^{-\beta\phi_{12}} - 1 = f_{12}, \tag{5.17}$$

and we obtain

$$W = \sum_k \frac{\beta_k}{k+1} n^k. \tag{5.18}$$

5.6. Ring Diagram Approximation for a Classical Electron Gas

For the Coulomb potential

$$\phi(r) = \frac{e^2}{r}, \tag{6.1}$$

we now evaluate the equation of state in the *ring diagram approximation*. A number $s+1$ electrons forming an interaction ring (Fig. 5.5) contribute to W

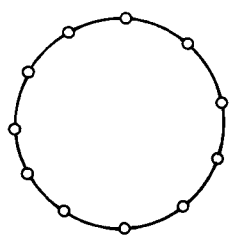


Fig. 5.5. A ring diagram.

in the form

$$\langle \phi_{12}\phi_{23}\cdots\phi_{s1}\rangle_c = V^{-(s-1)} \int \phi_{12}\phi_{23}\cdots\phi_{s1}\, d\mathbf{r}_2\, d\mathbf{r}_3 \cdots d\mathbf{r}_s.$$

The evaluation is easily performed by a Faltung (convolution) method. The Fourier transform of Eq. (6.1) is

$$u(q) = \int \phi(r) \exp i\mathbf{q}\cdot\mathbf{r}\, d\mathbf{r} = \frac{4\pi e^2}{q^2}. \tag{6.2}$$

This result can be obtained by introducing a convergence factor $e^{-\varepsilon r}$ and by letting ε be zero after integration. Using this Fourier transform, and also

$$\frac{1}{(2\pi)^3} \int \exp(i\mathbf{q}\cdot\mathbf{r})\, d\mathbf{r} = \delta(\mathbf{q}), \tag{6.3}$$

we find

$$\langle \phi_{12}\phi_{23}\cdots\phi_{s1}\rangle = \frac{1}{(2\pi)^3} \int [u(q)]^s\, d\mathbf{q}. \tag{6.4}$$

Equation (6.4) is divergent. However, if we proceed further to the evaluation of physical quantities, we get convergent results. As a matter of fact, we find for W of Eq. (5.13)

$$W_R = \frac{1}{(2\pi)^3} \sum_{s=2}^{\infty} \frac{1}{2} \frac{(s-1)!}{s!} \left\{ n^{s-1} \int [-\beta u(q)]^s\, d\mathbf{q} \right\}, \tag{6.5}$$

where $(s-1)!/2(s!)$ is the weight factor for a ring of s members. We interchange the summation with the integration and obtain the result

$$W_R = \frac{1}{4\pi^2 n} \int \left[\frac{\kappa^2}{q^2} - \log\left(1 + \frac{\kappa^2}{q^2}\right) \right] q^2\, dq, \tag{6.6}$$

where

$$\kappa = (4\pi\beta n e^2)^{1/2} \tag{6.7}$$

is the *Debye–Hückel screening constant*.

Using Eq. (6.6) in Eq. (5.14) and taking the limit of small κ, we arrive at

$$\frac{p}{kT} = n\left(1 - \frac{\pi^{1/2}}{3} e^3 \beta^{3/2} n^{1/2}\right). \tag{6.8}$$

The second term corrects the ideal gas law. It is determined by the ratio of the Coulomb energy at the average particle distance to the thermal energy.

In the evaluation of the pair distribution function in the *chain diagram* approximation, one encounters the following type of integral:

$$P_n(r_{1,n+2}) = \int \cdots \int f_{12} f_{23} \cdots f_{n+1,n+2}\, d\mathbf{r}_2 \cdots d\mathbf{r}_{n+1}. \tag{6.9}$$

The integration can be performed by a Faltung method. In fact, using the Fourier transform

$$f(\mathbf{r}) = \frac{1}{(2\pi)^3} \int (\exp i\mathbf{k}\cdot\mathbf{r}) \tilde{f}(k)\, d\mathbf{k}, \tag{6.10}$$

we find for $r = r_{1,n+2}$

$$P_n(r_{1,n+2}) = \frac{1}{(2\pi)^3} \int [\tilde{f}(k)]^{n+1} \exp(i\mathbf{k} \cdot \mathbf{r})\, d\mathbf{k}. \tag{6.11}$$

The reader might want to try to derive the Debye–Hückel-type distribution function

$$\rho_2 = n^2\left(1 - \varepsilon \frac{e^{-\kappa r}}{\kappa r}\right) \qquad (\varepsilon = \beta e^2 \kappa) \tag{6.12}$$

for a system of Coulomb forces.

5.7. Theory of Condensation

The theory of condensation is one of the most important unsolved problems in statistical mechanics. In 1938, Mayer (*5*) tried to interpret the onset of condensation as the point where the function represented by $\sum z^l b_l$ becomes nonanalytic as z is increased from zero. As one can see from the relation between z and n given by Eq. (2.7), to increase z means to increase n. The relation is one-to-one, at least for small values of n and z.

The Mayer theory does not prove the existence of the condensation point. Similar considerations have been forwarded by Born and Fuchs, Kahn and Uhlenbeck, and others. However, none of them actually "proved" the existence of condensation. In 1952, Yang and Lee proposed an entirely different theory. Although their theory also does not prove the existence of condensation, we discuss it briefly (*8*).

Consider a finite system of hard-sphere particles. Because the volume is finite, one can introduce only a finite number of particles into the system. Therefore, the grand partition function Ξ must terminate at a maximum possible particle number M. As a result, we obtain a power series

$$\Xi(T, v; z) = \sum_{N=0}^{M} z^N Q_N, \tag{7.1}$$

where Q_N is the configurational partition function for an N-particle system. Since Q_N is positive, Ξ does not vanish except for complex zeros z_k. We can write (since $Q_0 = 1$)

$$\Xi(T, V; z) = \prod_{k=1}^{M} \left(1 - \frac{z}{z_k}\right). \tag{7.2}$$

Equation (7.1) shows that Ξ increases monotonically with z, and so does $\log \Xi$. Here, z represents effectively the density and $\log \Xi$ the pressure.

Therefore, for a finite volume, there will be no condensation. We have to consider the possibility that in the limit $V \to \infty$ some complex zero may move to the positive part of the real axis.

For this limit, $V \to \infty$, it can be proved for that part of the positive half of the real axis that is free from zeros of Ξ that $\lim_{V\to\infty} \log \Xi$ exists and is a continuous analytic function. Its derivative, i.e., the density, is also analytic for the same part.

In such a region, we can use the expansion

$$\log\left(1 - \frac{z}{z_k}\right) = -\left[\frac{z}{z_k} + \frac{1}{2}\left(\frac{z}{z_k}\right)^2 + \frac{1}{3}\left(\frac{z}{z_k}\right)^2 + \cdots\right], \tag{7.3}$$

obtaining

$$\frac{p}{kT} = \lim_{V\to\infty} \sum b_j(V, T) z^j, \tag{7.4}$$

where

$$b_j(V, T) = -\frac{1}{Vj} \sum_{k=1}^{M} \left[\frac{1}{z_k(V, T)}\right]^j, \qquad j \geqslant 1. \tag{7.5}$$

Equation (7.4) is simply the cluster expansion of the pressure.

Zeros of Ξ can be expected on the positive axis only in the *thermodynamic limit*:

$$V \to \infty, \qquad N \to \infty, \qquad \frac{N}{V} = n \quad \text{(finite)}.$$

Combining the contributions to $\log \Xi$ from a zero of Ξ, $z_k = r_k e^{i\theta_k}$, and from its complex conjugate, $z_k^* = r_k e^{-i\theta_k}$, we find

$$\log\left(1 - \frac{z}{z_k}\right) + \log\left(1 - \frac{z}{z_k^*}\right) = 2 \log \frac{[z^2 - 2zr_k(\cos\theta_k) + r_k^2]^{1/2}}{r_k}$$
$$= -V(\phi_k(z) - \phi_k(0)) \tag{7.6}$$

where a new function $\phi_k(z)$ is introduced. The equation of state is given by

$$-2p\beta = -2\left(\frac{1}{V} \log \Xi\right) = \sum_k^M [\phi_k(z) - \phi_k(0)]. \tag{7.7}$$

Let us imagine that $\phi_k(z)$ is an electric potential. The corresponding electric field $E_k(z)$ is given by $-\partial\phi_k/\partial z$. The field $E_k(z)$ is related to the density by

$$\frac{2n}{z} = \frac{2}{z}\left[\frac{1}{V}\frac{\partial(\log \Xi)}{\partial(\log z)}\right] = \sum_k E_k(z). \tag{7.8}$$

The relation between Eqs. (7.7) and (7.8) is due merely to the nature of the grand partition function. Yet it becomes interesting if we remember that the electrostatic potential at a distance r from a line charge of density σ is of the form

$$\phi_k(z) = -2\sigma \log r. \tag{7.9}$$

Because Ξ is real, if a zero z_1 of Ξ approaches the real axis, its conjugate complex $z_1{}^*$ also has to approach the same spot. When the temperature is changed, the zeros of Ξ move around in the complex plane. The highest temperature T_c above which no zero falls on the real axis must be the critical temperature. As $T \to T_c$, two real zeros, z_1 and z_2, could approach each other, one of them corresponding to the gas–liquid and the other to the liquid–solid phase transition.

In Figs. 5.6 and 5.7, a line $P'P$ is perpendicular to a plane at P. Let us

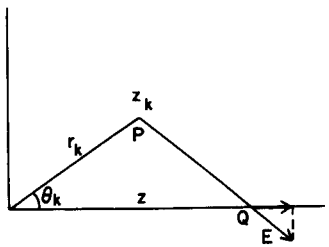


Fig. 5.6. An electrostatic analog of the grand partition function.

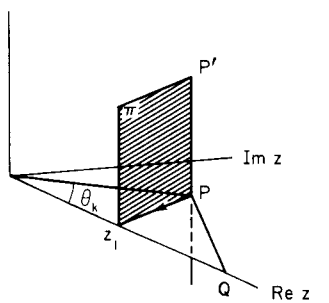


Fig. 5.7. A continuous distribution of charges on a sheet (zeros of Ξ approach the real axis).

imagine that $P'P$ represents a line charge and that the x and y axes of the plane correspond respectively to the real and imaginary axes of the complex plane that we have considered. The point P has circular coordinates (r_k, θ_k). The potential at a point Q of coordinates $(z, 0)$ due to the line charge is

$$\phi_k = -2\sigma \log[z^2 - 2zr_k(\cos\theta_k) + r_k{}^2]^{1/2}. \tag{7.10}$$

We find that, if the charge density σ is $1/V$, this expression reduces to Eq. (7.6). In this way, we can set up an electrostatic analog of the condensation problem.

If, in the limit $V \to \infty$, zeros of Ξ approach the positive real axis at z_1, we will have a continuous charge distribution on a two-dimensional sheet as is shown in Fig. 5.7. According to electrostatistics, the potential (i.e., p/kT) is continuous across the plane, but there is a discontinuity of the electric field of an amount

$$E(z_1+) - E(z_1-) = 4\pi\sigma(z_1). \tag{7.11}$$

In other words, one expects a first-order phase change, as is observed in gas–liquid condensation.

Figure 5.8 represents the isotherms of xenon near the critical temperature,

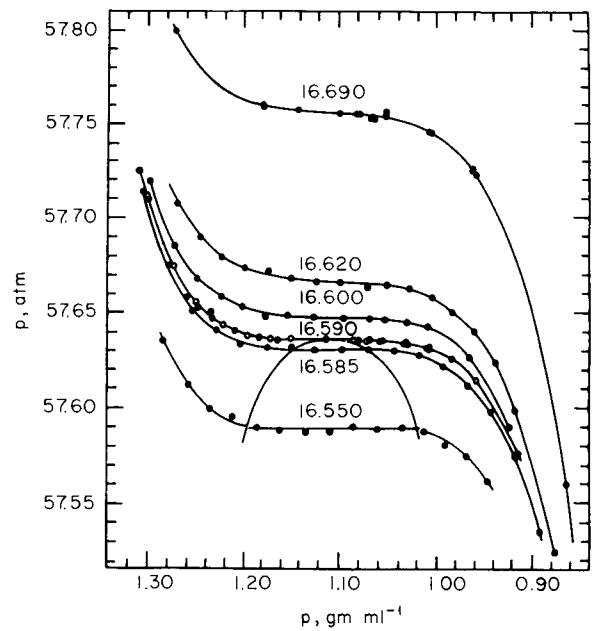


Fig. 5.8. Isotherms of xenon near the critical temperature. The numbers on the curves represent the temperature in degrees Celsius. [After Habgood and Schneider (*9*). Reproduced by permission of the National Research Council of Canada from the *Canadian Journal of Chemistry* **32**, 98–112 (1954).]

obtained by Habgood and Schneider (*9*). As we shall discuss in Chapter 9, the gas–liquid coexistence curve shown by the solid curve follows a simple power law. From the curves, we see that, at the critical temperature,

$$\left(\frac{\partial p}{\partial \rho}\right)_{T_c} = 0, \qquad \left(\frac{\partial^2 p}{\partial \rho^2}\right)_{T_c} = 0.$$

5.8. Polarizable Gases

We shall calculate the dielectric susceptibility and the dielectric constant of a polarizable gas in a homogeneous external field. For simplicity, we assume that the molecules are spherical and do not have permanent dipole moments but only induced dipoles.

The dielectric constant ε is obtained by calculating the average dipole moment of a unit volume, which is the polarization. The polarization $\mathbf{P}$ is related to the macroscopic field $\mathbf{E}$ by

$$\mathbf{P} = \frac{\varepsilon - 1}{4\pi}\mathbf{E}. \tag{8.1}$$

The macroscopic field is the average field acting on the molecules and differs from the externally imposed field, because of the polarization effect. If we assume that the molecules are small dielectric spheres, we have

$$\mathbf{E}_0 = \mathbf{E} + \tfrac{4}{3}\pi\mathbf{P},$$

that is, (8.2)

$$\mathbf{E}_0 = \tfrac{1}{3}(\varepsilon + 2)\mathbf{E},$$

where $\mathbf{E}_0$ is the external field. Thus, combining Eqs. (8.1) and (8.2), we obtain

$$\frac{\varepsilon - 1}{\varepsilon + 2} = \frac{4\pi}{3}\frac{P}{E_0}. \tag{8.3}$$

The right-hand side may be calculated by using statistical mechanics because the electric polarization $\mathbf{P}$ is the average dipole moment per unit volume.

The total field acting on the ith molecule is the external field $\mathbf{E}_0$ plus the field produced by the dipoles in the other molecules. The latter field is of the form $-\sum \mathbf{T}_{ik} \cdot \mathbf{p}_k$, where $\mathbf{T}_{ik}$ is the dipole–dipole interaction tensor defined by

$$\begin{aligned}\mathbf{T}_{ik} &= -\nabla\nabla\left(\frac{1}{r_{ik}}\right)\\ &= \frac{1}{r_{ik}^3}\left(\mathbf{1} - \frac{3\mathbf{r}_{ik}\mathbf{r}_{ik}}{r_{ik}^2}\right).\end{aligned} \tag{8.4}$$

It is assumed that the dipole moment in the ith molecule is proportional to the total field:

$$\mathbf{p}_i = \alpha\left(\mathbf{E}_0 - \sum_k \mathbf{T}_{ik} \cdot \mathbf{p}_k\right). \tag{8.5}$$

The proportionality constant α is the polarizability. We have N such equations, where N is the number of molecules. The formal solution of such equations is given by

$$\mathbf{p} = (1 + \alpha\mathbf{T})^{-1} \cdot \alpha\mathbf{E}_0, \tag{8.6}$$

where $\mathbf{p}$ is a vector with components $\mathbf{p}_i$ and $\mathbf{T}$ is a tensor with elements $\mathbf{T}_{ik}$, its diagonal elements being zero. The subscripts i and k run from 1 to N.

In the absence of the external field, the interaction energy is given by Eq. (1.2). To polarize the molecules, we need the "charging" energy:

$$\begin{aligned}V_c &= \sum_i \int_0^p \left(\mathbf{E}_0 - \sum_k \mathbf{T}_{ik} \cdot \mathbf{p}_k\right) \cdot d\mathbf{p}_i\\ &= \sum_i \frac{\mathbf{p}_i \cdot \mathbf{p}_i}{2\alpha}.\end{aligned} \tag{8.7}$$

The dipoles thus induced interact with the external field, giving the orientation energy

$$V_0 = -\sum_i \mathbf{p}_i \cdot \mathbf{E}_0. \tag{8.8}$$

Finally, the dipoles interact with each other with the energy

$$V_d = \tfrac{1}{2}\alpha \sum \mathbf{p}_i \cdot \mathbf{T}_{ik} \cdot \mathbf{p}_k. \tag{8.9}$$

Adding these energies, we get the total interaction Hamiltonian of the following form:

$$\begin{aligned}H &= H_1 + H_2\\ H_1 &= \sum_{i<k} \phi_{ij}, \qquad H_2 = -\tfrac{1}{2}\alpha E_0^2(\mathbf{1} + \alpha\mathbf{T})^{-1}_{zz},\end{aligned} \tag{8.10}$$

where the second term H_2 represents the effect due to the external field, assumed to be in the z direction.

The polarizability is very small, and we expand H_2 in powers of α:

$$H_2 = -\tfrac{1}{2}\alpha E_0^2[1 - \alpha T + (\alpha T)^2 - \cdots]_{zz}. \tag{8.11}$$

In this expansion, the lth term contains a product of l matrices in the N-dimensional molecular space and we shall take the zz elements of the matrices in the brackets.

Each element in the matrix $\mathbf{T}$ represents a dipole interaction of two particles. We see that the elements of $\mathbf{T}^l$ represent interactions that may be illustrated by diagrams having l bonds. The number of particles in these diagrams may vary from 2 to N, the total number of molecules in the system. The series for H_2 corresponds to the expansion in the number of bonds. However, each term in the series can be decomposed into terms according to

the number of particles. Therefore, the series for H_2 may be rearranged in such a way that the summation is first performed over all bonds of molecular clusters of a given size (number of molecules) and then over all sizes of clusters.

This alteration of the order of summations introduces an effective dipole interaction tensor $\mathbf{F}_{ij}$ for two particles, $\mathbf{F}_{ijk}$ for three particles, and so on. In terms of these functions, H_2 may be rewritten as follows:

$$\beta H_2 = -x(1 + \sum \mathbf{F}_{ij} + \sum\sum \mathbf{F}_{ijk} + \cdots)_{zz}, \tag{8.12}$$

where $\beta = 1/kT$, and x is a dimensionless quantity defined by

$$x = \tfrac{1}{2}\beta\alpha E_0{}^2. \tag{8.13}$$

To evaluate the integral of the configurational partition function given by Eq. (1.1), we introduce the two-body, three-body, and higher-order Mayer-type f-functions:

$$\begin{aligned} \exp[-\beta\phi_{ij} + x(\mathbf{F}_{ij})_{zz}] &= 1 + f_{ij}, \\ \exp[x(\mathbf{F}_{ijk})_{zz}] &= 1 + f_{ijk}. \end{aligned} \tag{8.14}$$

In constrast with f_{ij}, the three-body and higher f-functions are dependent only on the coupling between the induced dipoles and are characteristic of polarization effects.

Expanding the product of f-functions in the integrand of Q_N as in Mayer's theory and constructing the grand partition function Ξ, we obtain (*10*)

$$\log \Xi = V \sum b_s e^{sx} z^s. \tag{8.15}$$

This equation means that $\log \Xi$ is a sum of connected diagrams represented by the b_s.

For spherical molecules in a homogeneous field, the polarization $\mathbf{P}$ is in the direction of $\mathbf{E}_0$. Since H_2 given by Eq. (8.10) is equal to $\sum \mathbf{E}_0 \cdot \mathbf{p}_i/2$, we can evaluate the average dipole moments by differentiating Eq. (8.15) with respect to the parameter x, keeping z constant. In the limit of vanishing field, we arrive at the following result:

$$\begin{aligned} \frac{P}{E_0} &= \frac{\alpha}{V}\left[\frac{\partial(\log \Xi)}{\partial x}\right]_{x=0} \\ &= n\alpha + \alpha \sum_{s=2} \left(\frac{\partial b_s}{\partial x}\right)_{x=0} z^s, \end{aligned} \tag{8.16}$$

where n is the density in the limit of vanishing field:

$$n = \frac{1}{V}\frac{\partial(\log \Xi)}{\partial(\log z)} = \sum b_s s z^s. \tag{8.17}$$

z in Eq. (8.16) is to be determined from Eq. (8.17).

Using Eq. (8.16), we find that the left-hand side of Eq. (8.3) can be expressed as follows

$$\frac{\varepsilon - 1}{\varepsilon + 2} = \frac{4\pi}{3}\alpha n[1 + S(n, T)], \tag{8.18}$$

where S is a function of the density n and temperature T. In particular, for small density, we have the next approximate formula for S:

$$S = n\left[\tfrac{1}{2}\int e^{-\beta\phi}(F_{12})_{zz}\, d\mathbf{r}_{12}\right], \tag{8.19}$$

where

$$\begin{aligned} (F_{12})_{zz} &= \left(\frac{-2\alpha\mathbf{T}_{12}}{1 + \alpha\mathbf{T}_{12}}\right)_{zz} \\ &= 2\sum_{n=0}^{\infty} \frac{(-\alpha)^{n+1}}{r^{3(n+1)}}\left[1 - \frac{z^2}{r^2} + (-2)^{n+1}\frac{z^2}{r^2}\right] \end{aligned} \tag{8.20}$$

is an effective dipole interaction tensor between two particles. S is obtained by taking the statistical average of this effective dipole interaction tensor.

For $S = 0$, Eq. (8.18) reduces to the Clausius–Mossotti formula. The function S represents a correction to the Clausius–Mossotti formula. When the density is small, Eqs. (8.18) and (8.19) suggest that the plot of the quantity

$$\frac{3(\varepsilon - 1)}{(\varepsilon + 2)4\pi n\alpha}$$

against the density n gives a straight line. If we use the molar density n_0, we obtain

$$\frac{\varepsilon - 1}{(\varepsilon + 2)n_0} = A + Bn_0. \tag{8.21}$$

Figure 5.9 illustrates the data on methane obtained by Johnston *et al.* (*11*).

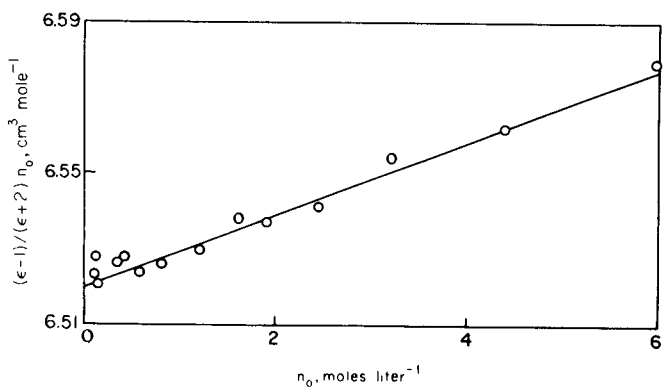


Fig. 5.9. Deviations from the Clausius–Mossotti formula. [After Johnston *et al.* (*11*).]

The ordinate represents the left side of Eq. (8.21). Such a deviation from the Clausius–Mossotti function has been observed on some other gases, including inert gases. Their results are summarized in Table 5.2.

TABLE 5.2

Values of A and B of the Clausius–Mossotti Function[a]

Gas	T(°K)	A (cm^3 mole^{-1})	B (cm^6 mole^{-2})
He	296	0.522	−0.5
A	242	4.138	0.6
	296	4.145	1.6
	306	4.140	1.0
N_2	242	4.389	4.2
	296	4.381	2.0
	306	4.384	1.8
	344	4.382	0.0
	344	4.380	1.5
CH_4	242	6.518	9.0
	315	6.543	7.3

[a] Note that, in this table, the value of A for He is approximately $\frac{1}{10}$ of those for other gases and that its B is negative.

Problem

Calculate the dielectric constant of an ideal gas composed of molecule shaving a permanent dipole moment μ. The dipole–dipole interactions may be neglected.

5.9. Bounds of the Free Energy

The Gibbs inequality (5.3) of Chapter 2 introduced in connection with the H-theorem may be used to derive an upper and a lower bound for the configurational free energy, which is related to the configurational partition function (*12*). The Gibbs inequality states, for arbitrary distribution functions $f(x)$ and $g(x)$ of the same norm, that

$$\int f(x) \log f(x)\, dx \geqslant \int f(x) \log g(x)\, dx. \tag{9.1}$$

Let us consider a system for which the Hamiltonian and the free energy split such that

$$H = H_0 + H_1, \qquad F = F_0 + F_1,$$

and let us define two distribution functions by

$$f(x) = e^{\beta(F_0 - H_0(x))}, \qquad g(x) = e^{\beta(F - H(x))}, \tag{9.2}$$

where x is a representative phase-space variable that can represent many variables and can even be discrete. Both distribution functions are normalized to 1. Actually, the free energies F_0 and F_1 are chosen such that

$$\int f(x)\, dx = \int g(x)\, dx = 1. \tag{9.3}$$

Thus,

$$e^{-\beta F_1} = \frac{\int e^{-\beta(H_0 + H_1)}\, dx}{\int e^{-\beta H_0}\, dx} = \langle e^{-\beta H_1} \rangle_0. \tag{9.4}$$

The value F_1 given by (9.4) may be expressed in terms of the Thiele semiinvariants C_s as

$$-\beta F_1 = \sum C_s \beta^s. \tag{9.5}$$

where

$$C_1 = \langle -H_1 \rangle_0,$$
$$C_2 = \tfrac{1}{2}(\langle H_1^2 \rangle_0 - \langle H_1 \rangle_0^2),$$
$$C_3 = \frac{1}{3!}\{\langle -H_1^3 \rangle_0 - 3\langle -H_1 \rangle \langle H_1^2 \rangle_0 - 2\langle H_1 \rangle_0^3\}, \quad \text{etc.}$$

with

$$\langle H_1^s \rangle_0 = \frac{\int e^{-\beta H_0} H_1^s\, dx}{\int e^{-\beta H_0}\, dx}. \tag{9.6}$$

The Gibbs inequality (9.1) gives

$$F_0 - \langle H_0 \rangle \geqslant \int (F - H) e^{\beta(F_0 - H_0)}\, dx, \tag{9.7}$$

or

$$\langle H_1 \rangle_0 \geqslant F_1. \tag{9.8}$$

In terms of the invariants, this inequality assumes the form

$$\frac{\beta}{2}(\langle H_1^2 \rangle_0 - \langle H_1 \rangle_0^2) \geqslant \frac{\beta^2}{3!}\{\langle H_1^3 \rangle - 3\langle H_1 \rangle_0 \langle H_1^2 \rangle_0 + 2\langle H_1 \rangle_0^3\}$$
$$- \frac{\beta^3}{4!}\{\langle H_1^4 \rangle_0 - 4\langle H_1 \rangle_0 \langle H_1^3 \rangle_0 - 3\langle H_1^2 \rangle_0^2$$
$$+ 12\langle H_1 \rangle_0^2 \langle H_1^2 \rangle_0 - 6\langle H_1 \rangle_0^4\} + \cdots. \tag{9.9}$$

We note that the Gibbs inequality is valid for arbitrary distribution functions of the same norm. Thus, we may use it for the case where f and g are interchanged:

$$f = e^{\beta(F-H)}; \qquad g = e^{\beta(F_0-H_0)}. \tag{9.10}$$

We then find

$$F_1 \geqslant \langle H_1 \rangle, \tag{9.11}$$

where

$$\langle H_1 \rangle = \frac{\int H_1 e^{-\beta H} dx}{\int e^{-\beta H} dx}. \tag{9.12}$$

If H_0 and H_1 do not have common variables as in the case of kinetic energy and potential energy, this expression reduces to

$$\langle H_1 \rangle = \int H_1 e^{\beta(F_1-H_1)}\, dx. \tag{9.13}$$

Using these equations in the Gibbs inequality and from the case where f and g are interchanged, we obtain (*12*)

$$\langle H_1 \rangle_0 \geqslant F_1 \geqslant \langle H_1 \rangle. \tag{9.14}$$

It is straightforward to extend our consideration to quantal systems. For this purpose, we introduce the grand potential Ω by

$$e^{-\beta\Omega} = \operatorname{tr} e^{-\beta(H-\mu N)} \tag{9.15}$$

and the normalized distribution function by

$$f = e^{\beta(\Omega - H + \mu N)}. \tag{9.16}$$

We split Ω formally as follows:

$$\Omega = \Omega_0 + \Omega_1. \tag{9.17}$$

As we shall discuss later, the grand potential is obtained from

$$e^{-\beta\Omega} = \operatorname{tr}[e^{-\beta(H_0-\mu N)} U(\beta)], \tag{9.18}$$

that is

$$\begin{aligned} e^{-\beta(\Omega-\Omega_0)} &= \langle U(\beta) \rangle_0 \\ &= 1 + \sum \frac{(-)^n}{n!} \int_0^\beta \langle P[H_1(x_1) \cdots H(x_n)]\, dx_1 \cdots dx_n \rangle_0 . \end{aligned} \tag{9.19}$$

Here, P is the chronological operator. If we use the notation of the second quantization (see Chapters 14 and 15) and express the Hamiltonian by

$$\begin{aligned} H_0 &= \sum \varepsilon_r a_r^* a_r \\ H_1 &= \frac{1}{2V} \sum_{\substack{rs \\ r's'}} (rs|\phi|r's') a_r^* a_s^* a_{s'} a_{r'}, \end{aligned} \tag{9.20}$$

we obtain

$$\begin{aligned} e^{-\beta(\Omega-\Omega_0)} = \sum_{n=1}^{\infty} \frac{(-)^n}{n!(2V)^n} \sum \prod_{i=1}^{n} (r_i s_i|\phi|r_i' s_i') \\ \times \int_0^\beta \langle P[a_{r_1}^*(x_1) a_{s_1}^*(x_1) a_{s_1'}(x_1) a_{r_1'}(x_1) \cdots] \rangle \prod dx_i . \end{aligned} \tag{9.21}$$

Introducing

$$\begin{aligned} \langle H_1 \rangle_0 &= \frac{\operatorname{tr} H_1 e^{-\beta(H_0-\mu N)}}{\operatorname{tr} e^{-\beta(H_0-\mu N)}}, \\ \langle H_1 \rangle &= \frac{\operatorname{tr} H_1 e^{-\beta(H-\mu N)}}{\operatorname{tr} e^{-\beta(H-\mu N)}}, \end{aligned} \tag{9.22}$$

we obtain

$$\langle H_1 \rangle_0 \geqslant \Omega_1 \geqslant \langle H_1 \rangle. \tag{9.23}$$

In particular, if

$$\langle H_1 \rangle_0 = 0, \tag{9.24}$$

we find

$$\Omega_D \geqslant \Omega, \tag{9.25}$$

where the subscript D stands for diagonal. This result is known as the *Peierls inequality* (*13*, *14*).

Problems

1. For

$$g(x) = \prod_i \left(\frac{\lambda}{\pi}\right)^{1/2} \exp(-\lambda x_i^2),$$

apply the inequality

$$\int f(x) \log f(x)\, dx \geqslant \int f(x) \log g(x)\, dx,$$

and show that the entropy

$$S = -k \int f(x) \log f(x)\, dx$$

is a maximum as a function of λ if

$$2\pi\langle x_i^2 \rangle = 1,$$

that is, if the equipartition law holds.

2. Show that, for the irreducible series of an imperfect gas, the following inequality holds:

$$\sum \frac{1}{k+1}\left(\frac{\beta_k}{\beta}\right)v^{-k} \leqslant \sum_k \frac{1}{k+1}\left(\frac{\partial\beta_k}{\partial\beta}\right)v^{-k},$$

where v is the specific volume.

5.10.* Cluster Expansions for Binary Mixtures

Let us consider a binary mixture, using the subscripts 1 for the solvent and 2 for the solute. The grand partition function is given by (*15–17*)

$$\Xi(T, V; z_1, z_2) = \sum_{N_1, N_2} Q_{N_1, N_2} z_1^{N_1} z_2^{N_2}. \tag{10.1}$$

and that for the pure solvent by

$$\Xi_0(T, V; z_1, 0) = \sum_{N_1} Q_{N_1, 0} z_1^{N_1}. \tag{10.2}$$

Here,

$$Q_{N_1, N_2} = \frac{1}{N_1!\, N_2!} \int e^{-\beta\Phi\{N_1, N_2\}}\, d\{N_1 + N_2\}, \tag{10.3}$$

and $\{N\}$ represents the coordinates of N particles. The osmotic pressure π is defined by

$$\pi = kT \frac{\partial[\log(\Xi/\Xi_0)]}{\partial V}. \tag{10.4}$$

By definition, it is clear that the ratio Ξ/Ξ_0 can be expanded in the form

$$\frac{\Xi}{\Xi_0} = \sum_{N_2} \frac{z_2^{*N_2}}{N_2!} \int e^{-\beta W\{N_2\}} d\{N_2\} \tag{10.5}$$

$$= \sum_{N_2} \frac{z_2^{*N_2}}{N_2!} \int F\{N_2\}\, d\{N_2\}, \tag{10.6}$$

where $z_2^* = z_2/z_2^0$, and z_2^0/kTv_1 is known as the Henry law constant. The probability distribution of solute particles in the presence of the solvent is $F\{N_2\}$. Inserting Eq. (10.6) into Eq. (10.4), we find in the limit of infinite dilution

$$\frac{\pi}{kT} = \sum_l b_l z_2^l, \tag{10.7}$$

where b_l is given by

$$b_l = \frac{1}{Vl!} \int g_l\{l\}\, d\{l\} \tag{10.8}$$

with the correlation functions $g_l\{l\}$ defined by

$$F_N\{N\} = \sum_{m_l} \prod_l g_l^{m_l}\{l\}. \tag{10.9}$$

Here, the summation of products corresponds to all the subdivisions of an integer N such that $N = \sum l m_l$. In particular,

$$\begin{aligned} g_2\{2\} &= F\{2\} - F\{1\}F\{1\}, \\ g_3\{3\} &= F\{3\} - \sum F\{1\}F\{2\} + 2F\{1\}F\{1\}F\{1\}. \\ &\vdots \end{aligned} \tag{10.10}$$

Differentiating Eq. (10.7) with respect to $c_2 = \langle N_2 \rangle / V$, we find

$$\frac{1}{kT}\left(\frac{\partial\pi}{\partial c_2}\right)_{T, v, z_1} = \sum_l l b_l z_2^l \left[\frac{\partial(\log z_2)}{\partial c_2}\right]_{T, v, z_1}. \tag{10.11}$$

We recall that the quantity $[\partial\langle N_2 \rangle/\partial(\log z_2)]_{T, v}$ represents the fluctuation $\langle N_2^2 \rangle - \langle N_2 \rangle^2$. Thus, we find

$$\frac{1}{kT}\left(\frac{\partial\pi}{\partial c_2}\right)_{T, v, z_1} = \frac{\langle N_2 \rangle}{\langle N_2^2 \rangle - \langle N_2 \rangle^2}. \tag{10.12}$$

On the other hand, the normalization conditions of $F\{N\}$ give

$$\int [F\{2\} - F\{1\}F\{1\}]\, d\{2\} = \langle N_2(N_2 - 1) \rangle - \langle N_2 \rangle^2. \tag{10.13}$$

Thus, we arrive at

$$\frac{1}{kT}\left(\frac{\partial\pi}{\partial c_2}\right)_{T, v, z_1} = \frac{1}{1 + c_2 G_{22}}, \tag{10.14}$$

where

$$G_{22} = \frac{V}{\langle N_2 \rangle^2} \int [F\{2\} - F\{1\}F\{1\}]\, d\{2\} \tag{10.15}$$

The distribution functions may be used to express other thermodynamic quantities. For instance, it can be shown that the chemical potential of the solvent for low molar concentration $x_2 = \langle N_2\rangle/(\langle N_1\rangle + \langle N_2\rangle)$ is

$$\frac{\mu_1 - \mu_1^0}{kT} = (\log x_1) + f_2 x_2^2 + \cdots, \tag{10.16}$$

where

$$f_2 = \frac{G_{11}^0 + G_{22}^0 - 2G_{12}^0}{2\overline{v_1^0}}. \tag{10.17}$$

Here, the G_{ij} are defined by integrals similar to Eq. (10.15) for pairs i, j. The superscript zero represents the zero-concentration limit and $\overline{v_1^0}$ is the partial volume of the solvent. In Eq. (10.16) the first term corresponds to an *ideal* (random) mixing. The second term is an interaction term that appears in a *regular solution* theory.

In the case of polymer solutions, expansions in the volume concentration ϕ, which is given by

$$\phi = \frac{\overline{v_2^0} c_2}{\overline{v_1^0} c_1 + \overline{v_2^0} c_2} \qquad \left[c_i = \frac{\langle N_i\rangle}{V}\right] \tag{10.18}$$

are more convenient. For instance, the expression corresponding to Eq. (10.16) is

$$\frac{\mu_1 - \mu_1^0}{kT} = -\frac{\phi}{N} + \frac{\phi^2}{N}\left(\frac{1}{2N^2} - 1 + \frac{1}{2}\frac{G_{22}^0}{N\overline{v_1^0}}\right) + O(\phi^3), \tag{10.19}$$

where $N = \overline{v_2^0}/\overline{v_1^0}$.

A similar free-energy expression has been obtained based on a lattice model. The ratio N of the solute partial molar volume to the solvent partial molar volume characterizes Eq. (10.19). The ϕ^2 interaction was first obtained experimentally by Van Laar.

Problems

1. Show that, for a two-component system of N_1 molecules of type 1 and N_2 molecules of type 2, the equation of state is approximately given by

$$\frac{pV}{(N_1 + N_2)kT} = 1 + \frac{1}{V}(x_1^2 b_{11} + 2x_1 x_2 b_{12} + x_2^2 b_{22}),$$

where $x_i = N_i/(N_1 + N_2)$ and the b_{ij} are the second virial coefficients for pairs of molecules i and j.

2. Calculate the second virial coefficient for the square-well potential

$$\phi(r) = \begin{cases} \infty, & r < a \\ -\phi_0, & a \leqslant r \leqslant b. \\ 0, & r > b \end{cases}$$

3. Show that the second and the third virial coefficients of a hard-sphere gas are given by

(*i*) two dimensions;

$$B = \frac{\pi}{2}a^2; \qquad C = B^2\left(\frac{4}{3} - \frac{\sqrt{3}}{\pi}\right),$$

(*ii*) three dimensions:

$$B = \tfrac{2}{3}\pi a^3; \qquad C = \tfrac{5}{8}B^2.$$

4. Using a Fourier transform method in summing the series

$$\phi_{\text{eff}} = \phi(r_{12}) - \beta n \int \phi_{13}\phi_{32}\, d\mathbf{r}_3 + (\beta n)^2 \int \phi_{13}\phi_{34}\phi_{42}\, d\mathbf{r}_3\, d\mathbf{r}_4 - \cdots,$$

derive the expression

$$\phi_{\text{eff}} = \frac{e^2}{r} e^{-\kappa r}.$$

for $\phi = e^2/r$.

5. Solve the Poisson–Boltzmann equation

$$\Delta\phi = -4\pi e^2[\delta(r) - n_0 e^{-\phi_{\text{eff}}/kT}]$$

by expanding the effective potential ϕ_{eff} as follows:

$$\phi_{\text{eff}} = \phi_0 + n\phi_1 + n^2\phi_2.$$

6. Find the Cauchy–Schwartz inequality in the following form:

$$\log \operatorname{tr}(e^{-\beta H_0} e^{-\beta H_1}) \leqslant \tfrac{1}{2}\log(\operatorname{tr} e^{-2\beta H_0}) + \tfrac{1}{2}\log(\operatorname{tr} e^{-2\beta H_1}).$$

7.* Investigate the inequality

$$\operatorname{tr} e^{-\beta(H_0 + H_1)} \leqslant \operatorname{tr}(e^{-\beta H_0} e^{-\beta H_1}),$$

where H_0 and H_1 are operators.

GENERAL REFERENCES

H. D. Ursell, *Proc. Camb. Philos. Soc.* **23**, 685 (1927); J. E. Mayer, **5**, 67 (1937).
E. E. Salpeter, *Ann. Phys.* **5**, 183 (1958); M. S. Green, *J. Math. Phys.* **1**, 391 (1960); C. DeDominicis, *J. Math. Phys.* **3**, 983 (1962); **4**, 255 (1963).

H. L. Frisch and J. L. Lebowitz, eds, "The Equilibrium Theory of Classical Fluids." Benjamin, New York, 1964.
T. L. Hill, "Statistical Mechanics." McGraw-Hill, New York, 1956.

REFERENCES

1. E. A. Guggenheim, *J. Chem. Phys.* **13**, 253 (1945).
2. A. Isihara, *J. Chem. Phys.* **19**, 397 (1951).
3. A. Isihara and T. Hayashida, *J. Phys. Soc. Japan* **6**, 40, 46 (1951).
4. E. Beth and G. E Uhlenbeck, *Physica* **4**, 915 (1937).
5. J. E. Mayer and M. G. Mayer, "Statistical Mechanics." McGraw-Hill, New York, 1940.
6. K. Husimi, *J. Chem. Phys.* **18**, 682 (1950).
7. J. E. Mayer and E. W. Montroll, *J. Chem. Phys.* **9**, 626 (1941).
8. C. N. Yang and T. D. Lee, *Phys. Rev.* **87**, 404, 410 (1952).
9. H. W. Habgood and W. G. Schneider, *Can. J. Chem.* **32**, 98 (1954).
10. A. Isihara, *J. Chem. Phys.* **38**, 2437 (1963); *Physica* **30**, 265 (1964).
11. D. R. Johnston, G. J. Oudeman, and R. H. Cole, *J. Chem. Phys.* **33**, 1310 (1960).
12. A. Isihara, *J. Phys. A (Proc. Phys. Soc.) Ser.* 2, **1**, 539 (1968).
13. M. D. Girardeau, *J. Chem. Phys.* **41**, 2945 (1964).
14. R. B. Griffith, *J. Math. Phys.* **5**, 1215 (1964).
15. W. G. McMillan and J. E. Mayer, *J. Chem Phys.* **13**, 276 (1945).
16. J. G. Kirkwood and F. P. Buff, *J. Chem. Phys.* **19**, 774 (1951).
17. A. Isihara and E. Guth, *Adv. Poly. Sci.* **5**, 233 (1967).

6

Distribution Functions

In statistical mechanics, many distribution functions are used, depending on our purpose. In this chapter, we shall be concerned mainly with the spatial distribution functions of classical systems. These distribution functions are introduced to describe both equilibrium and nonequilibrium phenomena.

The theory of the distribution functions of classical systems made remarkable strides around 1946 through the efforts of Bogoliubov (4), Born and Green, Kirkwood (1), Yvon, and others (2, 3, 5–7). We shall first consider time-dependent cases, following these authors. Then, we shall discuss the treatment of scattering phenomena in a fluid using the pair distribution function. We then treat the determination of the pair distribution function for a fluid in equilibrium. For quantum distribution functions, the reader should refer to Chapter 11.

6.1. Reduced Liouville Equation and Boltzmann Equation

For a homogeneous fluid composed of N molecules in a volume V, we introduce the probability $f^{(N)}(\mathbf{r}^N, \mathbf{p}^N; t)\, d\mathbf{r}^N\, d\mathbf{p}^N$ of finding the molecules in the specified domain $d\mathbf{r}^N\, d\mathbf{p}^N$ about the positions $\mathbf{r}^N = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$ and the momenta $\mathbf{p}^N = (\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N)$. The total distribution function $f^{(N)}$ satisfies the Liouville equation (1.7) of Chapter 2. It is assumed that the system is described by the Hamiltonian of Eq. (3.1) of Chapter 3. Thus,

$$\frac{df^{(N)}}{dt} = \frac{\partial f^{(N)}}{\partial t} + \sum_{i=1}^{N} \left(\frac{\mathbf{p}_i}{m} \cdot \frac{\partial}{\partial \mathbf{r}_i} f^{(N)} + \frac{\partial}{\partial \mathbf{p}_i} \cdot \mathbf{F}_i f^{(N)} \right) = 0, \tag{1.1}$$

where $\mathbf{F}_i$ is the force acting on the ith particle.

We define the reduced distribution function $f^{(l)}(\mathbf{r}^l, \mathbf{p}^l; t)$ by

$$f^{(l)}(\mathbf{r}^l, \mathbf{p}^l; t) = \frac{1}{(N-l)!} \int f^{(N)}(\mathbf{r}^N, \mathbf{p}^N; t)\, d\mathbf{r}_{l+1}\, d\mathbf{p}_{l+1} \cdots d\mathbf{r}_N\, d\mathbf{p}_N \tag{1.2}$$

and the reduced distribution function of l particles in coordinate space by

$$\rho_l(\mathbf{r}^l; t) = \int f^{(l)}(\mathbf{r}^l, \mathbf{p}^l; t)\, d\mathbf{p}^l. \tag{1.3}$$

In particular, $f^{(0)} = 1$ and ρ_1 is the number density. The normalization of these distribution functions has been defined such that

$$\begin{aligned} \int f^{(N)}\, d\mathbf{r}^N\, d\mathbf{p}^N &= N! \\ \int f^{(l)}\, d\mathbf{r}^l\, d\mathbf{p}^l &= N(N-1)\cdots(N-l+1). \end{aligned} \tag{1.4}$$

The normalization of $f^{(l)}$ may be understood if we recall that the distribution function $f^{(l)}$ does not state which l out of N particles are in the specified positions. Thus, if $f^{(l)}$ is integrated over the entire phase space, we should get $N(N-1)\cdots(N-l+1)$, which is the number of ways of selecting l indistinguishable objects out of N. From Eqs. (1.4), it follows that

$$\int f^{(l+1)}\, d\mathbf{r}^{l+1}\, d\mathbf{p}^{l+1} = (N-l) f^{(l)}. \tag{1.5}$$

Both $f^{(l)}(\mathbf{r}, \mathbf{p}; t)$ and $f^{(l)}(\mathbf{r} - [1/m]\mathbf{p}\, dt, \mathbf{p}; t - dt)$ give ρ_l upon integrating over the velocity space. Using this property, we can derive the equation of continuity:

$$\frac{\partial \rho_l}{\partial t} + \sum_{i=1}^{l} \frac{\partial}{\partial \mathbf{r}_i} \cdot (\rho_l \mathbf{u}_i^{(l)}) = 0, \tag{1.6}$$

where $\partial/\partial \mathbf{r}_i$ = grad, and $\mathbf{u}_i^{(l)}$ is the average velocity, defined by

$$\mathbf{u}_i^{(l)} = \frac{1}{\rho_l} \int f^{(l)} \frac{\mathbf{p}_i}{m} \prod d\mathbf{p}_i. \tag{1.7}$$

The differential equation for $f^{(l)}$ can be derived by taking into consideration the acceleration of the molecules due to their interactions. Assuming that the interaction potentials are dependent only upon relative molecular distances, we obtain

$$\begin{aligned} &\frac{\partial f^{(l)}}{\partial t} + \sum_{i=1}^{l} \frac{\partial f^{(l)}}{\partial \mathbf{r}_i} \cdot \frac{\mathbf{p}_i}{m} - \sum_{i \neq j}^{l} \frac{\partial f^{(l)}}{\partial \mathbf{p}_i} \cdot \frac{\partial \phi_{ij}}{\partial \mathbf{r}_i} \\ &\quad = \iint \sum_{i=1}^{l} \frac{\partial \phi_{i,l+1}}{\partial \mathbf{r}_i} \cdot \frac{\partial f^{(l+1)}}{\partial \mathbf{p}_i}\, d\mathbf{r}_{l+1}\, d\mathbf{p}_{l+1}. \end{aligned} \tag{1.8}$$

This equation may also be derived by reducing the Liouville equation for $f^{(N)}$ to that for $f^{(l)}$ by integrating over the coordinates of $N - l$ particles. The left-hand side of Eq. (1.8) represents the change of $f^{(l)}$ in the phase space of l molecules. The right-hand side is the change of $f^{(l)}$ due to the interaction of the l particles with the rest of the particles in the total system and is called the *collision term*. The existence of this term shows that $f^{(l)}$ does not satisfy the same Liouville equation as $f^{(N)}$.

According to Eq. (1.8), the differential equation for $f^{(1)}$ involves $f^{(2)}$ on the right-hand side, that for $f^{(2)}$ depends on $f^{(3)}$, and so on. We have a hierarchy of equations—the BBGKY hierarchy, so called after Bogoliubov, Born, Green, Kirkwood, and Yvon. The determination of $f^{(1)}$ can be made only when the hierarchy is cut off somewhere.

For this purpose, let us consider a dilute gas, where the three-body interactions may be neglected in the differential equation for $f^{(2)}$. Using the relative coordinate $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ and assuming that $f^{(2)}$ does not depend on time explicitly, we obtain from Eq. (1.8)

$$\left(\frac{\mathbf{p}_2}{m} - \frac{\mathbf{p}_1}{m} \right) \cdot \frac{\partial f^{(2)}}{\partial \mathbf{r}} = \frac{\partial \phi}{\partial \mathbf{r}} \cdot \left(\frac{\partial f^{(2)}}{\partial \mathbf{p}_2} - \frac{\partial f^{(2)}}{\partial \mathbf{p}_1} \right). \tag{1.9}$$

This approximate result may be used on the right-hand side of Eq. (1.8) for $f^{(1)}$. We find

$$\begin{aligned} &\frac{\partial f^{(1)}}{\partial t} + \frac{\mathbf{p}_1}{m} \cdot \frac{\partial f^{(1)}}{\partial \mathbf{r}_1} + \mathbf{F} \cdot \frac{\partial f^{(1)}}{\partial \mathbf{p}_1} \\ &\quad = \iint \left[-\frac{\partial \phi}{\partial \mathbf{r}} \cdot \frac{\partial f^{(2)}}{\partial \mathbf{p}_2} + \left(\frac{\mathbf{p}_2}{m} - \frac{\mathbf{p}_1}{m} \right) \cdot \frac{\partial f^{(2)}}{\partial \mathbf{r}} \right] d\mathbf{r}_2\, d\mathbf{p}_2, \end{aligned} \tag{1.10}$$

where on the left-hand side a term due to an external force $\mathbf{F}$ has been added. The first term on the right-side vanishes because of Gauss's theorem, and the second term can be simplified as follows.

We have been using potentials which are functions of relative distances. Therefore, a collision of two particles takes place in a plane. As is illustrated in Fig. 6.1, let us consider the case where molecule 2 approaches molecule 1 with relative velocity $\mathbf{g} = (\mathbf{p}_2/m) - (\mathbf{p}_1/m)$, changes its direction of motion by an angle χ upon collision, and leaves with relative velocity $\mathbf{g}'$. The

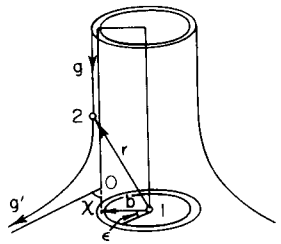


Fig. 6.1. Two-particle collision.

absolute magnitude of $\mathbf{g}$ is denoted by $|\mathbf{g}| = g$. In Fig. 6.1 O is the point of closest approach. We denote by $\mathbf{b}$ the perpendicular which is the vector drawn from 1 to the asymptote. The integration over the coordinates of molecule 2 which appeared on the right-hand side of Eq. (1.10) may be carried out first along the orbit and then over all possible $\mathbf{b}$'s. The volume integral of the gradient of $f^{(2)}$ may be converted into a surface integral, since $\mathbf{g}$ is in the direction of the tangent to the orbit. Thus, the second integral on the right-hand side of Eq. (1.10) is transformed as follows:

$$\iint [f^{(2)}(+) - f^{(2)}(-)]\mathbf{g}\, d\mathbf{b}\, d\mathbf{p}_2,$$

where $+$ and $-$ correspond, respectively, to after $(\mathbf{r}\cdot\mathbf{g} > 0)$ and before $(\mathbf{r}\cdot\mathbf{g} < 0)$ the collision. We replace $f^{(2)}$ for large particle separations by the product of two $f^{(1)}$'s. We obtain

$$\frac{\partial f_1}{\partial t} + \frac{\mathbf{p}}{m}\cdot\frac{\partial f_1}{\partial \mathbf{r}} + \mathbf{F}\cdot\frac{\partial f_1}{\partial \mathbf{p}} = \iint (f_1' f_2' - f_1 f_2)\mathbf{g}\, d\mathbf{b}\, d\mathbf{p}_2, \tag{1.11}$$

where the quantities with primes mean that their arguments correspond to after the collision, and where $f^{(1)}(\mathbf{r}_i, \mathbf{p}_i; t)$ is written as $f_i^{(1)}$ for $i = 1, 2$. Equation (1.11) may be called the Boltzmann equation, although the original Boltzmann equation applies to the velocity distribution function. The above derivation of the Boltzmann equation is approximate, and more refined theories have been developed. As in Chapter 1, the detailed balance condition

$$f_1' f_2' = f_1 f_2 \tag{1.12}$$

corresponds to equilibrium.

It is important to recognize in the above derivation that the integration along the orbit has been performed in a *specified direction* such that f' or f_1' corresponds to after the collision. By this process, we have introduced a *time arrow*, i.e., irreversibility. The simplifications such as low density and no triplet encounter introduced in the derivation have nothing to do with irreversibility.

6.2. Stress Tensor in Nonequilibrium Fluids

In terms of the distribution function $f^{(N)}$, the average of a function $G(\mathbf{r}^N, \mathbf{p}^N)$ can be defined by

$$\langle G\rangle = \frac{1}{N!}\int f^{(N)}\, G\, d\mathbf{r}^N\, d\mathbf{p}^N. \tag{2.1}$$

Generally, the quantity on the left-hand side depends on time. Its time change is determined by

$$\frac{\partial\langle G\rangle}{\partial t} = \frac{1}{N!}\int \frac{\partial f^{(N)}}{\partial t}\, G\, d\mathbf{r}^N\, d\mathbf{p}^N$$

$$= -\frac{1}{N!}\sum_{i=1}^{N}\int G\left(\frac{\mathbf{p}_i}{m}\cdot\frac{\partial}{\partial \mathbf{r}_i} f^{(N)} + \frac{\partial}{\partial \mathbf{p}_i}\cdot \mathbf{F}_i f^{(N)}\right) d\mathbf{r}^N\, d\mathbf{p}^N.$$

By Gauss's theorem, this equation is reduced to

$$\frac{\partial\langle G\rangle}{\partial t} = \sum_{i=1}^{N}\left\langle \frac{\mathbf{p}_i}{m}\cdot\frac{\partial}{\partial \mathbf{r}_i} G + \mathbf{F}_i\cdot\frac{\partial}{\partial \mathbf{p}_i} G\right\rangle. \tag{2.2}$$

For instance, if $\rho = \sum_i m\,\delta(\mathbf{r}_i - \mathbf{r})$ is chosen as G, we obtain the equation of continuity

$$\frac{\partial \rho}{\partial t} + \frac{\partial}{\partial \mathbf{r}}\cdot\rho\mathbf{u} = 0,$$

$$\rho\mathbf{u} = \left\langle \sum_i \mathbf{p}_i\,\delta(\mathbf{r}_i - \mathbf{r})\right\rangle. \tag{2.3}$$

The equation of continuity is what Eq. (1.6) states for the case $l = 1$.

In many cases, it is more convenient to use the reduced distribution functions than to use $f^{(N)}$. We shall show in this and the next sections that Eq. (1.8) is very useful to obtain various expressions for average physical quantities related to $f^{(l)}$.

We shall derive the equation of motion of fluids starting with Eq. (1.8). For $l = 1$, we have

$$\frac{\partial f^{(1)}}{\partial t} + \frac{\partial f^{(1)}}{\partial \mathbf{r}_1}\cdot\frac{\mathbf{p}_1}{m} = \iint \frac{\partial \phi_{12}}{\partial \mathbf{r}_1}\cdot\frac{\partial f^{(2)}}{\partial \mathbf{p}_1}\, d\mathbf{r}_2\, d\mathbf{p}_2. \tag{2.4}$$

Multiplying Eq. (2.4) by the vector $\mathbf{p}_1/m = \mathbf{v}_1$, integrating over $\mathbf{p}_1$, and applying Gauss's theorem we find

$$\frac{\partial}{\partial t}(n\mathbf{u}) + \frac{\partial}{\partial \mathbf{r}_1}\cdot\int f^{(1)}\mathbf{v}_1\mathbf{v}_1\, d\mathbf{p}_1 = -\frac{1}{m}\int \rho_2 \frac{\partial \phi}{\partial \mathbf{r}_1}\, d\mathbf{r}_2 \qquad (\phi \equiv \phi_{12}). \tag{2.5}$$

Similarly, Eq. (1.6) for $l = 1$ [or Eq. (2.3)] multiplied by $\mathbf{u}$ yields

$$\mathbf{u}\frac{\partial n}{\partial t} - n\mathbf{u}\cdot\frac{\partial}{\partial \mathbf{r}_1}\mathbf{u} + \frac{\partial}{\partial \mathbf{r}_1}\cdot(n\mathbf{u}\mathbf{u}) = 0. \tag{2.6}$$

We introduce the relative coordinate $\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1$ and the center-of-gravity coordinate $\mathbf{R} = \frac{1}{2}(\mathbf{r}_1 + \mathbf{r}_2)$, and expand $\rho_2(\mathbf{r}_1, \mathbf{r}_2) = \rho_2(\mathbf{r}, \mathbf{R})$ as follows:

$$\begin{aligned}\rho_2(\mathbf{r}, \mathbf{R}) &= \rho_2(\mathbf{r}, \mathbf{r}_1 + (\mathbf{R} - \mathbf{r}_1)) \\ &= \rho_2(\mathbf{r}, \mathbf{r}_1) + (\mathbf{R} - \mathbf{r}_1)\cdot\frac{\partial}{\partial \mathbf{r}_1}\rho_2(\mathbf{r}, \mathbf{r}_1) + \cdots.\end{aligned} \tag{2.7}$$

We may neglect higher-order terms, since ρ_2 varies slowly with the position $\mathbf{r}_1$. The first term, $\rho_2(\mathbf{r}, \mathbf{r}_1)$, of Eq. (2.7) does not contribute to the integral on the right-hand side of Eq. (2.5). By using the second term and subtracting Eq. (2.6) from Eq. (2.5), the following macroscopic equation of motion is obtained:

$$mn\frac{d}{dt}\mathbf{u} - \frac{\partial}{\partial \mathbf{r}_1}\cdot\boldsymbol{\sigma} = 0, \tag{2.8}$$

where

$$-\boldsymbol{\sigma} = m\int f^{(1)}(\mathbf{v}_1\mathbf{v}_1 - \mathbf{u}\mathbf{u})\, d\mathbf{p}_1 - \frac{1}{2}\int\frac{\mathbf{r}\mathbf{r}}{r}\rho_2\phi'(r)\, d\mathbf{r} \tag{2.9}$$

is the stress tensor. Later, we shall derive an analogous equation for equilibrium fluids.

In order to go further, it is necessary to know $f^{(1)}$ and ρ_2 for nonequilibrium systems. However, Eq. (1.8) is not useful as it stands; it is necessary to express the right-hand side in terms of $f^{(1)}$ in order to close the equation. The Boltzmann equation (1.11) is closed, but it is based on several assumptions. In the next section, we shall present another equation, which is considered to be better than the Boltzmann equation when applied to dense gases or liquids.

6.3. Viscosity Coefficient of Fluids

Let us investigate the equation for $f^{(1)}$. The right-hand side of Eq. (2.4) may be rewritten as

$$\begin{aligned}-\int\frac{\partial\phi}{\partial \mathbf{r}_1}\cdot\frac{\partial f^{(2)}}{\partial \mathbf{p}_1}\, d\mathbf{r}_2\, d\mathbf{p}_2 &= -\frac{\partial}{\partial \mathbf{p}_1}\cdot\int\frac{\partial\phi}{\partial \mathbf{r}_1}f^{(2)}\, d\mathbf{r}_2\, d\mathbf{p}_2 \\ &= m\frac{\partial}{\partial \mathbf{p}_1}\cdot\langle\mathbf{X}\rangle f^{(1)},\end{aligned} \tag{3.1}$$

where $m\langle\mathbf{X}\rangle$ is the average molecular force acting on any particle in the system.

When a Brownian particle moves in a fluid, its motion is described by the well-known Langevin equation:

$$\dot{\mathbf{p}} = -\zeta\left(\frac{\mathbf{p}}{m} - \mathbf{u}\right) + m\mathbf{K}, \tag{3.2}$$

where ζ is a frictional coefficient and $m\mathbf{K}$ is a newly added random thermal force. According to *Stokes's law*, $\zeta = 6\pi\eta a$ for a spherical particle of radius a moving in a fluid of viscosity η.

It is considered that $\langle\mathbf{X}\rangle$ balances the force on the right-hand side of Eq. (3.2) (*8*) such that

$$m\langle\mathbf{X}\rangle f^{(1)} = -\zeta(\mathbf{v} - \mathbf{u})f^{(1)} + m\mathbf{K}^{(1)}. \tag{3.3}$$

The random force $m\mathbf{K}$ has been introduced so that the phase point moves in and out of the momentum interval $(\mathbf{p}, \mathbf{p} + d\mathbf{p})$. Effectively, it causes a kind of diffusion in momentum space. Thus, we replace the term $\mathbf{K}f^{(1)}$ by $-D\partial f^{(1)}/\partial\mathbf{p}$. Here, D, the effective diffusion constant, can be determined by the condition that both sides of Eq. (3.3) vanish in equilibrium. Since the equilibrium distribution is of the form $\exp[-(\mathbf{p} - m\mathbf{u})^2/2mkT]$, we use $D = kT\zeta/m$ and put

$$m\langle\mathbf{X}\rangle f^{(1)} = -\zeta\left[(\mathbf{v} - \mathbf{u})f^{(1)} + kT\frac{\partial}{\partial\mathbf{p}}f^{(1)}\right]. \tag{3.4}$$

This is a generalized force-balancing equation.

Using Eq. (3.4), we now have a closed equation to determine $f^{(1)}$:

$$\frac{\partial f^{(1)}}{\partial t} + \frac{\mathbf{p}}{m}\cdot\frac{\partial}{\partial\mathbf{r}}f^{(1)} = \zeta\frac{\partial}{\partial\mathbf{p}}\cdot\left[(\mathbf{v} - \mathbf{u})f^{(1)} + kT\frac{\partial}{\partial\mathbf{p}}f^{(1)}\right]. \tag{3.5}$$

Kirkwood treated Eq. (3.4) in a more rigorous way and derived the above equation with a molecular expression for the frictional constant.

A similar differential equation may be derived for $f^{(2)}$,

$$\frac{\partial f^{(2)}}{\partial t} + \sum_{i=1}^{2}\mathbf{v}_i\cdot\frac{\partial f^{(2)}}{\partial\mathbf{r}_i} = \sum_{i=1}^{2}\zeta\frac{\partial}{\partial\mathbf{p}_i}\cdot\left[(\mathbf{v}_i - \mathbf{u})f^{(2)} + kT\frac{\partial}{\partial\mathbf{p}_i}f^{(2)}\right], \tag{3.6}$$

which will be used shortly.

Multiplying both sides of Eq. (3.5) by the tensor $(\mathbf{v} - \mathbf{u})(\mathbf{v} - \mathbf{u})$, integrating the result over momentum space, and using Green's theorem, we find a relation which connects the kinetic part of the stress tensor given by Eq. (2.9) with the frictional coefficient ζ. Retaining terms up to the first order in the velocity gradient, we obtain

$$m\int f^{(1)}(\mathbf{v}_1\mathbf{v}_1 - \mathbf{u}\mathbf{u})\, d\mathbf{p}_1 = nkT\mathbf{1} - \frac{nmkT}{\zeta}\left[\dot{\mathbf{s}} - \frac{1}{3}\left(\frac{\partial}{\partial\mathbf{r}}\cdot\mathbf{u}\right)\mathbf{1}\right], \tag{3.7}$$

where **1** is a unit tensor and $\dot{\mathbf{s}}$ is the symmetrized tensor representing the rate of strain. Its components are

$$\dot{s}_{ij} = \frac{1}{2}\left(\frac{\partial u_j}{\partial x_i} + \frac{\partial u_i}{\partial x_j}\right), \qquad i, j = x, y, z. \tag{3.8}$$

On the other hand, in viscous fluids, the stress tensor $\boldsymbol{\sigma}$ is expressed in terms of the coefficients of shear viscosity η and bulk viscosity κ as follows:

$$\boldsymbol{\sigma} = -\left[p + \left(\frac{2\eta}{3} - \kappa\right)\frac{\partial}{\partial \mathbf{r}} \cdot \mathbf{u}\right]\mathbf{1} + 2\eta\dot{\mathbf{s}}, \tag{3.9}$$

where p is the equilibrium pressure and **1** is a unit tensor.

Comparing the terms on the right-hand side of Eq. (3.7) with the corresponding terms in Eq. (3.9), we find the following momentum contribution to viscosity:

$$\eta_0 = \frac{nmkT}{2\zeta}. \tag{3.10}$$

In order to evaluate the contribution η_1 from the molecular forces to the coefficient of viscosity, the *pair correlation function* or the *radial distribution function* $g(\mathbf{r})$ is expanded to first order in the rate of shear and volume change:

$$g(r) = g_0^{(2)}(r)\left\{\left[1 + \frac{\zeta}{2kT}\left(\frac{\mathbf{r}_{12} \cdot \dot{\mathbf{s}} \cdot \mathbf{r}_{12}}{r_{12}^2} - \frac{1}{3}\operatorname{div}\mathbf{u}\right)\right]\Psi_2(r_{12}) + \frac{\zeta}{6kT}(\operatorname{div}\mathbf{u})\Psi_1(r_{12})\right\},$$

$$g(r) = \frac{\rho_2(r)}{n^2}, \tag{3.11}$$

where $g_0^{(2)}$ is the equilibrium radial distribution function and the Ψ_i are unknown functions. Note that the right-hand side is linear with respect to $\dot{\mathbf{s}}$ and div **u** and is independent of a choice of the coordinate system.

Using Eqs. (3.11) and (3.6), we are able to set up the differential equations for the determination of the unknown functions Ψ_i. These are

$$\frac{d}{dr}\left(r^2 g_0^{(2)}\frac{d\Psi_2}{dr}\right) - 6g_0^{(2)}\Psi_2 = r^3\frac{dg_0^{(2)}}{dr},$$

$$\frac{d}{dr}\left(r^2 g_0^{(2)}\frac{d\Psi_1}{dr}\right) = r^3\frac{dg_0^{(2)}}{dr} - 3r^2\frac{dg_0^{(2)}}{d(\log n)}. \tag{3.12}$$

The solutions of these equations are subject to the boundary conditions that the excess relative flow must vanish at infinity in addition to the requirement

of the normal behavior of the pair distribution function. The functions Ψ_i yield the transport coefficients due to interactions as follows:

$$\eta \cong \eta_1 = \frac{\pi\zeta n^2}{15kT}\int r^3\frac{d\phi}{dr}\Psi_2(r)g_0^{(2)}(r)\,dr,$$

$$\kappa = \frac{\pi\zeta n^2}{9kT}\int r^3\frac{d\phi}{dr}\Psi_1(r)g_0^{(2)}(r)\,dr, \tag{3.13}$$

where κ is the bulk viscosity.

The total shear viscosity coefficient is given by

$$\eta = \eta_0 + \eta_1. \tag{3.14}$$

These results are due to Kirkwood, Buff, and Green (*2*). Using an approximation for $g_0^{(2)}$, these investigators evaluated the shear viscosity of argon at the boiling point and found the value 0.73×10^{-3} P, which is of the order of the experimental value 2.39×10^{-3} P. The coefficient of thermal conductivity can be evaluated in a similar way.

6.4. Plasmas

So far, we have considered only uncharged particles. When the system is composed of charged particles and electromagnetic fields which are confined to a vessel of volume V, the total Hamiltonian is

$$\mathscr{H} = K + \Phi + H_f. \tag{4.1}$$

Here, K is the kinetic energy, Φ represents the Coulomb potential, and H_f is the Hamiltonian for the electromagnetic fields. These fields may be represented by an infinite number of harmonic oscillators because only the standing waves suitable to the boundary are allowed. We have

$$K = \sum_i \frac{1}{2m_i}\left(\mathbf{p}_i - \frac{e_i}{c}\mathbf{A}_i\right)^2,$$

$$\Phi = \sum_{i<j}\phi(r_{ij}),$$

$$H_f = \tfrac{1}{2}\sum_\lambda (p_\lambda^2 + \omega_\lambda^2 q_\lambda^2), \tag{4.2}$$

$$\omega_\lambda^2 = c^2 k_\lambda^2, \qquad \mathbf{k}_\lambda = \frac{2\pi}{V^{1/3}}\mathbf{n},$$

$$\mathbf{n} = (n_1, n_2, n_3) \qquad \text{(integers)}.$$

In the kinetic energy expression, $\mathbf{A}_i$ is the vector potential for the ith particle with mass m_i.

The distribution function for the entire system satisfies the Liouville equation. The distribution function of a single charged particle is obtained by integrating the total distribution function over the coordinates of all the fields and the other particles. It is interesting to express the resulting singlet distribution function in velocity-coordinate space rather than in momentum-coordinate space. The transformation from the latter to the former is done by using the relation between the momentum $\mathbf{p}_i$ and the velocity $\mathbf{v}_i$:

$$m_i \mathbf{v}_i = \mathbf{p}_i - \frac{e_i}{c}\mathbf{A}_i, \tag{4.3}$$

where the vector potential $\mathbf{A}_i$ is, in general, composed of two terms, one describing the externally imposed field and the other corresponding to the electromagnetic fields,

$$\mathbf{A}_i = \mathbf{A}_i^{\mathrm{e}} + \sum_{\lambda} q_\lambda(t)\mathbf{A}_\lambda(\mathbf{r}_i). \tag{4.4}$$

Using Eqs. (4.3) and (4.4), we obtain the following differential equation for the singlet distribution function $f^{(1)}$:

$$\frac{\partial f^{(1)}}{\partial t} + \mathbf{v}\cdot\frac{\partial f^{(1)}}{\partial \mathbf{r}} + \frac{e}{m}\mathbf{F}\cdot\frac{\partial f^{(1)}}{\partial \mathbf{v}} = \frac{1}{m} f^{(1)} \frac{\partial}{\partial \mathbf{v}}\cdot\left\langle\frac{\partial \phi}{\partial \mathbf{r}}\right\rangle, \tag{4.5}$$

where $\mathbf{F}$ is the Lorentz force defined by

$$\mathbf{F} = \mathbf{E} + \frac{1}{c}\mathbf{v}\times\mathbf{H}, \tag{4.6}$$

with

$$\mathbf{E} = -\frac{1}{c}\frac{\partial \mathbf{A}}{\partial t} - \left\langle\frac{\partial \phi}{\partial \mathbf{r}}\right\rangle, \qquad \mathbf{H} = \mathrm{rot}\,\mathbf{A}.$$

Here,

$$\left\langle\frac{\partial \phi}{\partial \mathbf{r}}\right\rangle = \frac{1}{f^{(1)}}\int \frac{\partial \phi}{\partial \mathbf{r}} f^{(2)}\, d\mathbf{v}_2\, d\mathbf{r}_2. \tag{4.7}$$

The average force $\langle -\partial\phi/\partial\mathbf{r}\rangle$ defined by this equation may not depend strongly on $\mathbf{v}$. In this case, we simplify Eq. (4.5) as follows:

$$\frac{\partial f^{(1)}}{\partial t} + \mathbf{v}\cdot\frac{\partial f^{(1)}}{\partial \mathbf{r}} + \frac{e}{m}\mathbf{F}\cdot\frac{\partial f^{(1)}}{\partial \mathbf{v}} = 0. \tag{4.8}$$

This equation holds if the pair distribution function $f^{(2)}$ is separable in velocity and coordinates or if it is a product of two singlet distribution functions.

6.5. Virial Equation of State

Let us start with the distribution function $f^{(N)}$ given by

$$Z_N f^{(N)} = e^{-\beta H} \tag{5.1}$$

for equilibrium fluids. Here, Z_N is the partition function of an N-particle system. We can integrate Z_N over the momentum space when the Hamiltonian is of the form

$$H = \sum_i \frac{p_i^2}{2m} + \sum_{i<j}\phi(r_{ij}). \tag{5.2}$$

The result of the integration is

$$Q_N \rho_N^{(N)} = e^{-\beta\Phi}, \tag{5.3}$$

where Q_N is the configurational partition function defined by Eq. (1.1) of Chapter 5. Further integration of this function over the coordinate space gives the reduced distribution function for l particles:

$$\rho_l^{(N)}(\mathbf{r}) = \frac{1}{Q_N(N-l)!}\int\cdots\int e^{-\beta\Phi}\prod_{i=l+1}^{N} d\mathbf{r}_i. \tag{5.4}$$

In particular, the pair distribution function (hereafter, p.d.f.) is

$$\rho_2^{(N)}(\mathbf{r}) = \frac{1}{Q_N(N-2)!}\int\cdots\int e^{-\beta\Phi}\prod_{i=3}^{N} d\mathbf{r}_i. \tag{5.5}$$

Using these distribution functions, we can evaluate average values of functions defined in the phase space. If a function of coordinates is to be averaged, $\rho^{(N)}$ may be used instead of $f^{(N)}$. In particular, if the function depends only on the coordinates of a small number of molecules, we can use Eq. (5.4). The difference between $f^{(N)}$ and $\rho^{(N)}$ is analogous to that between the partition function and the configurational partition function.

In Chapter 2, the statistical mechanical expressions for the entropy, free energy, and internal energy were presented. The internal energy given by Eq. (2.4) of Chapter 3 may be expressed as follows:

$$\frac{U}{NkT} = \frac{3}{2} + \frac{1}{2nkT}\int \phi(r)\rho_2(r) 4\pi r^2\, dr, \tag{5.6}$$

where the total potential energy is the sum of pair potentials as in Eq. (5.2).

The equation of state can be derived from this expression by differentiation with respect to the volume. Alternatively, we can use the formula $p = kT\partial(\log Q_N)/\partial V$ and Eqs. (5.3) and (5.5). In carrying out differentiation with

respect to the volume, we introduce dimensionless variables x_i such that the particle coordinates $\mathbf{r}_i$ are related to the volume by $\mathbf{r}_i = \mathbf{x}_i V^{1/3}$. We obtain

$$pV = NkT - \frac{V}{6}\int \rho_2(r) r \frac{\partial \phi}{\partial r} 4\pi r^2 \, dr. \tag{5.7}$$

Equation (5.7) is called the *virial equation of state* because it can be derived from Clausius's virial theorem. In Eqs. (5.6) and (5.7), the superscript N to the p.d.f. indicating the canonical ensemble is omitted. In fact, these equations which hold for the p.d.f., $\rho_2(r)$, can be defined in terms of the grand ensemble. The definition is

$$\rho_2(r) = \frac{1}{\Xi}\sum_{N=2} \frac{z^N}{(N-2)!}\int \cdots \int e^{-\beta\Phi} \prod_{i=3}^{N} d\mathbf{r}_i, \tag{5.8}$$

where $\Xi = \sum_N z^N Q_N$ is the grand partition function. The p.d.f. is normalized as follows:

$$\int \rho_2(r) \, d\mathbf{r} = \langle N(N-1) \rangle. \tag{5.9}$$

The Hamiltonian of the type of Eq. (5.2) may be written in the form

$$H = H_0 + \xi H_1, \tag{5.10}$$

where ξ is a coupling parameter. If the p.d.f. corresponding to this Hamiltonian is written as $\rho_2(r, \xi)$, one can show that the pressure p is given by

$$p - p_0 = -\tfrac{1}{2}\int_0^1 d\xi \int \phi(r)\rho_2(r, \xi) \, d\mathbf{r}, \tag{5.11}$$

where p_0 is the pressure corresponding to H_0.

Because of these relations, the p.d.f. is very important in the theory of fluids. Equation (5.11) may be used for the evaluation of the pressure of a quantum mechanical system.

Expressions similar to Eqs. (5.6) and (5.8) can be derived for surface quantities. For instance, the internal energy for a unit surface is approximately given by

$$u = -\tfrac{1}{2}\pi \int_0^\infty r^3 \phi(r)\rho_2(r) \, dr. \tag{5.12}$$

The surface tension γ is

$$\gamma = \tfrac{1}{8}\pi \int_0^\infty r^4 \phi'(r)\rho_2(r) \, dr, \tag{5.13}$$

where $\phi'(r)$ represents the derivative of $\phi(r)$.

6.6. Determination of Fluid Structure

The structure of fluids has been determined by X-ray scattering measurements. It is known that, structurally, liquids are closer to crystals than to gases. This is particularly true near the melting point. Since X-ray scattering is determined by interference, the p.d.f. is important for the evaluation of the scattering intensity.

The Thomson scattering formula

$$I_e = I_0 \frac{1}{R^2}\left(\frac{e^2}{mc^2}\right)^2 \frac{1 + \cos^2\theta}{2} \tag{6.1}$$

gives the scattering intensity at a distance R from a single electron with charge e and mass m and in the direction θ from the incident wave of strength I_0. The scattering of X-rays from fluids is determined by the interference of scattered waves from intraatomic and interatomic electrons. The former is expressed in terms of the so-called atomic structure factor f, which has been tabulated for many atoms. The latter is determined by the p.d.f. The combination of these two factors gives the scattering intensity in the following form:

$$I = NI_0 f^2(1 + i(s)), \tag{6.2}$$

where N is the total number of the particles and $i(s)$ is defined by

$$i(s) = \frac{1}{n}\int [\rho_2(r) - n^2] \frac{\sin sr}{sr} 4\pi r^2 \, dr. \tag{6.3}$$

Here, n is the number density, and

$$s = \frac{4\pi \sin(\theta/2)}{\lambda},$$

λ being the wavelength. It is convenient to introduce the radial distribution function $g(r)$ defined by

$$\rho_2 = n^2 g(r). \tag{6.4}$$

Equation (6.3) is rewritten as

$$i(s) = n\int_0^\infty [g(r) - 1] \frac{\sin sr}{sr} 4\pi r^2 \, dr. \tag{6.5}$$

Combining Eqs. (6.2) and (6.5), the scattering intensity can be calculated in terms of the distribution function. Actually, the structure of fluids is determined from the scattering intensity by a Fourier inversion formula as follows:

$$g(r) = 1 + \frac{1}{2\pi^2}\frac{1}{nr}\int_0^\infty si(s) \sin rs \, ds. \tag{6.6}$$

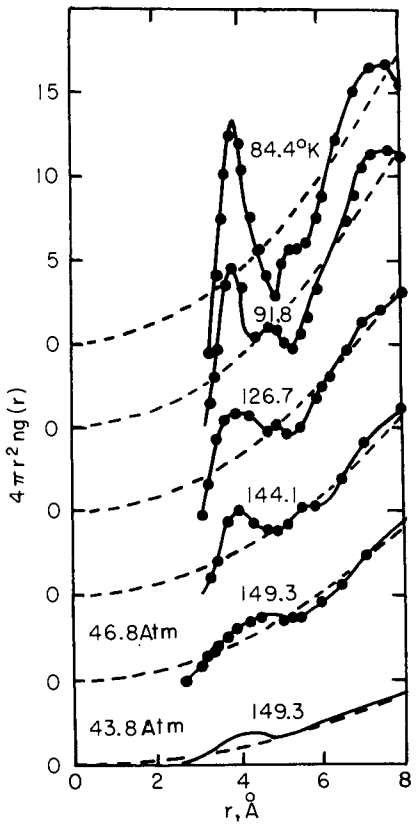


Fig. 6.2. The radial distribution functions of argon obtained by Gingrich (*9*).

Figure 6.2 shows $4\pi r^2 ng(r)$ of argon at various temperatures plotted against r (*9*). The quantity $4\pi r^2 ng(r)$ represents the number of particles in the spherical shell of unit width at a distance r from a particle. The first peak is seen to be higher at lower temperatures. Approximating the first peak by a Gaussian distribution, we can determine the "nearest-neighbor atoms." The number of nearest neighbors at 84.4°K lies between 10.2 and 10.9. This result indicates that the structure of liquid argon is between that of the fcc and the bcc lattices (*9*).

6.7. Critical Opalescence

As another example of making use of the p.d.f., we shall discuss *critical opalescence*. The scattering of light from a gas is usually not strong, but near the critical temperature it becomes strong due to large density fluctuations.

The connection between the fluctuations and the scattering intensity was first discussed by Ornstein and Zernike (*12a*).

Let us consider in a fluid a small volume V_1 which contains N_1 molecules. We may express N_1 in terms of a function $\Delta(\mathbf{r})$ which is 1 when $\mathbf{r}$ is in V_1 and 0 when it is outside. Since any one of the total N molecules can be in the volume V_1, we have

$$N_1 = \sum_{i=1}^{N} \Delta(\mathbf{r}_i); \qquad \Delta(r) = \begin{cases} 1 & \mathbf{r} \text{ in } V_1 \\ 0 & \mathbf{r} \text{ not in } V_1. \end{cases} \tag{7.1}$$

For N_1, we find the following averages:

$$\langle N_1 \rangle = nV_1 \tag{7.2}$$

and

$$\langle N_1^2 \rangle = \int_{V_1} \rho_1(\mathbf{r}_1)\, d\mathbf{r}_1 + \iint_{V_1} \rho_2(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_1\, d\mathbf{r}_2 . \tag{7.3}$$

The fluctuation of the number of molecules is expressed by

$$\frac{\langle N_1^2 \rangle - \langle N_1 \rangle^2}{\langle N_1 \rangle} = 1 + 4\pi n^{-1} \int_{V_1} [\rho_2(r) - n^2] r^2\, d\mathbf{r}. \tag{7.4}$$

According to the X-ray scattering formula (6.3), the right-hand side of Eq. (7.4) is equal to $1 + i(0) = I(0)$, which is proportional to the intensity of scattered light at zero angle:

$$I(0) = 1 + 4\pi n^{-1} \int_0^\infty [\rho_2(r) - n^2] r^2\, dr. \tag{7.5}$$

The fluctuation given by Eq. (7.4) can be related to a certain thermodynamic quantity. In order to find such a relation, we differentiate the chemical potential μ with respect to $\langle N \rangle$. For simplicity, we shall denote $\langle N \rangle$ by N whenever this abbreviation will not cause confusion. We start with

$$N\, d\mu = -S\, dT + V\, dp, \tag{7.6}$$

and differentiate this equation with respect to N at constant temperature. We obtain

$$\left(\frac{\partial \mu}{\partial N}\right)_{T,V} = \frac{V}{N}\left(\frac{\partial p}{\partial N}\right)_{T,V} = \frac{kT}{N}\left[\frac{\partial(\log \Xi)}{\partial N}\right]_{T,V}. \tag{7.7}$$

Differentiating this again, we get

$$\left(\frac{\partial^2 \mu}{\partial N^2}\right)_{T,V} = -\frac{kT}{N^2}\left[\frac{\partial(\log \Xi)}{\partial N}\right]_{T,V} + \frac{kT}{N}\left[\frac{\partial^2(\log \Xi)}{\partial N^2}\right]_{T,V}. \tag{7.8}$$

Thus, combining Eqs. (7.8) and (7.7), we obtain

$$\left[\frac{\partial^2(\log \Xi)}{\partial N^2}\right]_{T,V} = \frac{N}{kT}\left(\frac{\partial^2 \mu}{\partial N^2}\right)_{T,V} + \frac{1}{kT}\left(\frac{\partial \mu}{\partial N}\right)_{T,V}. \tag{7.9}$$

On the other hand, the left-hand side is equal to the following expression, as can be seen by changing the variable of differentiation:

$$\left[\frac{\partial^2(\log \Xi)}{\partial N^2}\right]_{T,V} = \left(\frac{1}{kT}\right)^2 \left[\frac{\partial^2(\log \Xi)}{\partial(\log \lambda)^2}\right]_{T,V} \left(\frac{\partial \mu}{\partial N}\right)^2_{T,V} + \frac{N}{kT}\left(\frac{\partial^2 \mu}{\partial N^2}\right)_{T,V}. \tag{7.10}$$

Equating Eqs. (7.9) and (7.10), we find

$$\left[\frac{\partial^2(\log \Xi)}{\partial(\log \lambda)^2}\right]_{T,V} = kT\left[\left(\frac{\partial \mu}{\partial N}\right)_{T,V}\right]^{-1}. \tag{7.11}$$

We note the thermodynamic equality:

$$\begin{aligned}\left(\frac{\partial \mu}{\partial N}\right)_{T,V} &= \left(\frac{\partial \mu}{\partial N}\right)_{T,p} + \left(\frac{\partial \mu}{\partial p}\right)_{T,N}\left(\frac{\partial p}{\partial N}\right)_{T,V} \\ &= -\left(\frac{\partial \mu}{\partial p}\right)_{T,N} \frac{\left(\frac{\partial V}{\partial N}\right)_{T,p}}{\left(\frac{\partial V}{\partial p}\right)_{N,T}}.\end{aligned} \tag{7.12}$$

Using Eq. (7.12) in Eq. (7.11), we obtain

$$\frac{\langle N^2\rangle - \langle N\rangle^2}{\langle N\rangle} = -\frac{NkT}{V^2}\left(\frac{\partial V}{\partial p}\right)_T. \tag{7.13}$$

The left-hand side coincides with the fluctuation given by Eq. (7.4). In other words, we have obtained

$$-\frac{NkT}{V^2}\left(\frac{\partial V}{\partial p}\right)_T = 1 + 4\pi n^{-1}\int_0^\infty [\rho_2(r) - n^2]r^2\, dr. \tag{7.14}$$

Combining this with the scattering intensity given by Eq. (7.5), we have

$$I(0) = -\frac{NkT}{V^2}\left(\frac{\partial V}{\partial p}\right)_T. \tag{7.15}$$

The quantity $-V^{-1}(\partial V/\partial p)_T$ is the *isothermal compressibility*. Denoting this by κ, we find

$$\kappa = \frac{1}{nkT}\left\{1 + 4\pi n^{-1}\int_0^\infty [\rho_2(r) - n^2]r^2\, dr\right\} \tag{7.16}$$

$$= (nkT)^{-1} I(0). \tag{7.17}$$

Therefore, the intensity of scattered light at zero angle is related to the compressibility. The compressibility of a fluid diverges at the critical temperature. This corresponds to critical opalescence.

To describe scattering of light due to density fluctuations in a fluid, it is convenient to express the number density at a position $\mathbf{r}$ as

$$n(r) = \left\langle \sum_i \delta(\mathbf{r} - \mathbf{r}_i) \right\rangle, \tag{7.18}$$

where δ represents a Dirac delta function and the bracket means a statistical average. The delta functions "pick up" particles whenever they come to the position $\mathbf{r}$. Equation (7.18) is analogous to Eq. (7.1).

Likewise, the p.d.f. can be expressed by

$$\rho_2(r) = V^{-1}\left\langle \sum_{i\neq j} \delta(\mathbf{r} - \mathbf{r}_i + \mathbf{r}_j) \right\rangle, \tag{7.19}$$

where $\mathbf{r}$ now corresponds to the relative vector between a pair of particles. The reader may confirm Eqs. (7.18) and (7.19) by checking the normalizations.

The *density–density correlation function* has been used in the theory of neutron scattering. It is defined by $\langle n(\mathbf{r})n(\mathbf{r}')\rangle$. From Eq. (7.19), we find

$$\langle n(\mathbf{r})n(\mathbf{r}')\rangle = n\delta(\mathbf{r} - \mathbf{r}') + \rho_2(|\mathbf{r} - \mathbf{r}'|). \tag{7.20}$$

6.8. Expansions of Distribution Functions

In Chapter 5, we introduced the cluster expansion of the grand partition function. Starting with this expansion, we are able to express various thermodynamic functions in cluster series and also in power series in density. Since thermodynamic functions can be expressed in terms of the distribution functions, cluster expansions may be expected to be possible for the distribution functions. In this section, we present the general cluster expansion theory (*10*) and then consider the density expansions for classical systems (*11*, *12*).

6.8.1. The Cluster Expansion

Let us introduce the function $W(\mathbf{r}^N)$ by

$$W(\mathbf{r}^N) = e^{-\Phi/kT}, \tag{8.1}$$

and define U-functions as follows:

$$\begin{aligned} W(\mathbf{r}^h) &= U_h(\mathbf{r}^h), \\ W(\mathbf{r}^h\mathbf{r}_k) &= U_{h+1}(\mathbf{r}^h\mathbf{r}_k) + U_h(\mathbf{r}^h)U_1(\mathbf{r}_k), \\ W(\mathbf{r}^h\mathbf{r}_k\mathbf{r}_l) &= U_{h+2}(\mathbf{r}^h\mathbf{r}_k\mathbf{r}_l) + \sum U_{h+1}(\mathbf{r}^h\mathbf{r}_k)U_1(\mathbf{r}_l) + U_h(\mathbf{r}^h)U_1(\mathbf{r}_k)U_1(\mathbf{r}_l), \end{aligned} \tag{8.2}$$

where $\mathbf{r}^h$ represents the set of coordinates of h particles and $\mathbf{r}_k$ is the set for the kth particle. Alternatively, these U-functions give the W-functions in the following way:

$$\begin{aligned} U_h(\mathbf{r}^h) &= W_h(\mathbf{r}^h), \\ U_{h+1}(\mathbf{r}^h\mathbf{r}_k) &= W_{h+1}(\mathbf{r}^h\mathbf{r}_k) - W_h(\mathbf{r}^h)W_1(\mathbf{r}_k). \end{aligned} \tag{8.3}$$

From Eq. (8.1), we have the relation

$$\begin{aligned} W(\mathbf{r}^N) = U_h(\mathbf{r}^h)W_{N-h}(\mathbf{r}_{h+1}\cdots\mathbf{r}_N) + \sum_k U_{h+1}(\mathbf{r}^h\mathbf{r}_k)W_{N-h-1} \\ + \sum_k\sum_l U_{h+2}(\mathbf{r}^h\mathbf{r}_k\mathbf{r}_l)W_{N-h-2} + \cdots. \end{aligned} \tag{8.4}$$

If we define the generalized "cluster integrals" by

$$b_l^{(h)}(\mathbf{r}^h) = \frac{1}{l!}\int U(\mathbf{r}^h\mathbf{r}_{h+1}\cdots\mathbf{r}_{h+l-1})\,d\mathbf{r}_{h+1}\cdots d\mathbf{r}_{h+l-1}, \tag{8.5}$$

we find

$$\rho_h(\mathbf{r}^h) = \sum_{l=1} l b_l^{(h)}\frac{Q_{N-h-l+1}}{Q_N}. \tag{8.6}$$

For large N, the ratio Q_{N-1}/Q_N is independent of N and is equal to the activity z. Thus, we arrive at the cluster expansion:

$$\rho_h(\mathbf{r}^h) = z^{h-1}\sum l b_l^{(h)}(\mathbf{r}^h)z^l. \tag{8.7}$$

In particular, for $h = 1$, Eq. (8.7) reduces to

$$n^{(1)} = n = \sum l b_l z^l, \tag{8.8}$$

and for $h = 2$ to

$$\rho_2(\mathbf{r}_{12}) = \sum l b_l^{(2)}(\mathbf{r}_{12})z^{l+1}. \tag{8.9}$$

Here, the first two b_l are given by

$$\begin{aligned} b_1^{(2)}(\mathbf{r}_{12}) &= W(\mathbf{r}_1, \mathbf{r}_2), \\ 2!\, b_2^{(2)}(\mathbf{r}_{12}) &= \int [W(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) - W(\mathbf{r}_1, \mathbf{r}_2)W(\mathbf{r}_3)]\,d\mathbf{r}_3. \end{aligned} \tag{8.10}$$

6.8.2. The Density Expansion

Let us expand $g(r) = \rho_2(r)/n^2$ in the density as follows:

$$g(r) = e^{-\beta\phi}\left[1 + \sum_{s=1} g_s(r)\, n^s\right]. \tag{8.11}$$

and try to prove that (*11*)

$$g_s(r) = \frac{1}{s!}\int\cdots\int\sum_2\prod f_{ij}\,d\mathbf{r}_3\cdots d\mathbf{r}_{s+2}, \tag{8.12}$$

where the f_{ij} are the Mayer f-functions. The summation and products in the integrand do not include f_{12}. The integral of g_s is characterized by the fact that it gives the irreducible integral for $s + 2$ particles upon integration over $\mathbf{r}_1$ and $\mathbf{r}_2$ if f_{12} is added.

We shall confirm Eq. (8.12) by showing that its substitution into Eq. (5.7) gives the virial expansion of Eq. (3.7) of Chapter 5. Thus, we start with

$$p = nkT\left[1 + \frac{1}{6}\int rf'(r)\sum_{s=0}^{\infty} g_s(r)n^{s+1}4\pi r^2\,dr\right], \tag{8.13}$$

where the summation includes the term corresponding to $g_0 = 1$. Let us define

$$I = \int\cdots\int_{(s)} rf'(r)\sum_2\prod f_{ij}\,d\mathbf{r}_3\cdots d\mathbf{r}_{s+2}\,d\mathbf{r}_2. \tag{8.14}$$

By introducing a dimensionless variable $x = r/L$, ($L^3 = V$), and integrating Eq. (8.14) over L by parts, Eq. (8.14) is transformed into

$$\begin{aligned} LI = \frac{2L^{3s+4}}{(s+1)(s+2)}\int\cdots\int_{(s+1)}\sum_1\prod f(r_{ij})\,d\mathbf{x}_2\cdots d\mathbf{x}_{s+2} \\ - \frac{(3s+4)2}{(s+1)(s+2)}\int_0^L\int\cdots\int_{(s+1)}\sum_1\prod f(r_{ij})L^{3s+3}\,d\mathbf{x}_2\cdots d\mathbf{x}_{s+2}\,dL. \end{aligned} \tag{8.15}$$

Here, $\sum_1$ means a summation over the products pertaining to irreducible cluster integrals involving $s + 1$ points. The factor $2/[(s + 1)(s + 2)]$ comes in because $\sum_2$ in $g_s(r)$ is symmetric in s particles, while $\sum_1$ is symmetric in $s + 2$ particles. Here, I and the products of f's can be considered to be independent of L.

It is easy to see that the last integrand is independent of L. Thus, upon integration, we find

$$I = -\frac{3(s+1)2}{(s+1)(s+2)}\int\cdots\int_{(s+1)}\sum_1\prod f(r_{ij})\,d\mathbf{r}_2\cdots d\mathbf{r}_{s+2}. \tag{8.16}$$

Substituting this into Eq. (8.13), we obtain

$$\begin{aligned} p &= nkT\left[1 - \sum\frac{n^{s+1}}{(s+2)s!}\int\cdots\int\sum_1\prod f(r_{ij})\,d\mathbf{r}_2\cdots d\mathbf{r}_{s+2}\right] \\ &= nkT\left(1 - \sum\frac{s}{s+1}n^s\beta_s\right), \end{aligned} \tag{8.17}$$

which is Eq. (3.7) of Chapter 5.

6.9.* Nodal Expansion

Equation (8.11) may be rewritten as follows:

$$g(r) = e^{-\beta\phi+S}, \tag{9.1}$$

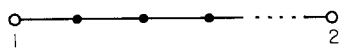
where

$$e^S = \sum_{s=0}^{\infty} g_s(r)n^s. \tag{9.2}$$

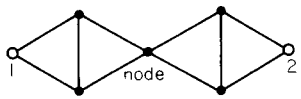
Note that the summation includes the term $s = 0$, g_0 being 1. The function $g_s(r)$ is given by those irreducible diagrams that involve a pair of particles at $\mathbf{r}_1$ and $\mathbf{r}_2$. We shall designate these two particles as 1 and 2 and represent them by open dots. The particles with coordinates to be integrated over will be drawn as darkened dots. Since only irreducible diagrams are under consideration, there will be no black dots with a single bond: All black dots have at least two bonds. Most generally, diagrams with systematic structures may be obtained by the following steps:

(*i*) *Formation of an "essentially series diagram."*

The term "essentially series" means either a single series connection of *f*-bonds (Fig. 6.3a) or a series connection of more complex graphs (Fig. 6.3b). The connecting points in the latter case are called *nodal points.*



(a)



(b)

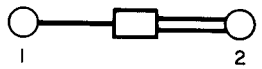
(c)

Fig. 6.3. "Essentially series" diagrams.

(*ii*) *Formation of a generalized chain.*

An essentially series diagram is represented by a single line called an effective bond. A generalized chain can be formed by a series connection of effective bonds and will be denoted by a heavy line (Fig. 6.3c). Nodal points, as the connecting elements, shall be represented by rectangles in such a general chain (Fig. 6.4). The terminals 1 and 2 are illustrated by open circles.

Fig. 6.4. A diagram with three effective lines.



(*iii*) *Parallel connections with effective lines.*

Many effective lines can connect the two terminals.

Let us explain these steps. An effective bond represents the total of diagrams that can be reduced to a chain if their internal structures are neglected. For instance, the diagramof Fig. 6.3a and 6.3b both belong to this category. A node is a point that must be always passed in order to go from 1 to 2 in a given diagram for $g_s(r)$.

We find at once that any number of independent effective lines exists between 1 and 2. Therefore, we decompose g_s, on the right-hand side of Eq. (9.2) as follows:

$$g_s(r)n^s = \sum_{m_l} \prod_l \frac{[\beta_l(r)n^l]^{m_l}}{m_l!}, \qquad \sum l m_l = s. \tag{9.3}$$

This decomposition is analogous to that in Eq. (2.2) or Eq. (3.3) of Chapter 5. Thus, we find

$$S = \sum \beta_l(r)n^l, \tag{9.4}$$

where the right-hand side is now a *summation over irreducible line diagrams* of arbitrary number of nodes.

We can further reduce the expression for S by introducing a topological consideration of the chain diagrams. In general, the diagrams contributing to S of Eq. (9.4) can be classified into two groups, one having nodes and the other without nodes. If we denote the contribution from the diagrams with nodes by N and that from the diagrams without nodes by E, then

$$S = N + E. \tag{9.5}$$

Let us consider diagrams with an arbitrary number of nodes and let the point 3 be the closest node to the point 1. Between 3 and 2, there can be any number of effective interaction lines (see Fig. 6.4) which altogether could

represent $G(r_{32}) = g(r_{32}) - 1$. Here, the term -1 is necessary in order to compensate the first term (which is 1) of the expansion of $g(r_{32})$.

There exists at least one (effective or true) interaction line between 3 and 2. Let us introduce X, defined as $G - N$:

$$G = X + N. \tag{9.6}$$

Combining this with Eqs. (9.1) and (9.6), we obtain

$$\begin{aligned} X &= G - N \\ &= e^{-\beta\phi}\, e^{S} - 1 - N, \\ G &= e^{-\beta\phi + S} - 1. \end{aligned} \tag{9.7}$$

Since $nG(r_{32})\, d\mathbf{r}_3$ is the probability of finding particle 3 in the volume element $d\mathbf{r}_3$ at a position $\mathbf{r}_{32}$ from particle 2, one obtains the integral equation

$$N(\mathbf{r}_{12}) = n \int X(\mathbf{r}_{13}) G(\mathbf{r}_{32})\, d\mathbf{r}_3 . \tag{9.8}$$

Thus, we obtain

$$G(\mathbf{r}_{12}) = X(\mathbf{r}_{12}) + n \int X(\mathbf{r}_{13}) G(\mathbf{r}_{32})\, d\mathbf{r}_3 . \tag{9.9}$$

This equation, which is exact, shall be called the *Ornstein–Zernike equation* (*12a*).

Unfortunately, the integral equation (9.9) is not closed: It is necessary to know $X(r)$ in order to solve the equation. For this purpose, several approximations have been proposed. We shall discuss some of them in the next section.

6.10.* HNC and PY Approximations

Recent developments concerning the integral equation for the pair correlation function are more or less based on the theories developed by Mayer (*7*), Born and Green (*5*), Kirkwood (*1*), and others (*2–6*). Therefore, it is natural to start with these classical theories. However, to save space, let us introduce two typical approximations first and then discuss the classical Born–Green theory.

6.10.1.* Hypernetted Chain (HNC) Approximation

This approximation is sometimes called the *convolution approximation* (*13–17*). It amounts to replacing $N(r)$ by $S(r)$:

$$N(r) \to S(r). \tag{10.1}$$

Then,

$$X(r) = g(r) - 1 - \log g(r) - \beta\phi(r), \tag{10.2}$$

and Eq. (9.8) becomes

$$\begin{aligned} \log g(r_{12}) + \beta\phi(r_{12}) = n \int & [g(r_{13}) - 1 - \log g(r_{13}) - \beta\phi(r_{13})] \\ & \times [g(r_{32}) - 1]\, d\mathbf{r}_3 . \end{aligned} \tag{10.3}$$

One sees from Eq. (9.5) that the *HNC approximation* assumes

$$E = 0. \tag{10.4}$$

By expanding $g(r)$ in linked cluster diagrams, one can find the nature of the approximation in detail. Actually, the name of this approximation is due to such a consideration. By numerical analyses, it has been found that the HNC approximation yields fairly good results at low temperature and at moderate density for attractive potentials.

6.10.2.* Percus–Yevick (PY) Approximation

This approximation replaces N as follows (*18*):

$$N(r) \to e^{S(r)} - 1. \tag{10.5}$$

Thus, $X(r)$ of Eq. (9.7) becomes

$$X(r) = (e^{-\beta\phi(r)} - 1) e^{S(r)} = g(r)(1 - e^{\beta\phi(r)}), \tag{10.6}$$

and the integral equation (9.8) is replaced by

$$N(r_{12}) = n \int [G(r_{13}) - N(r_{13})] G(r_{32})\, d\mathbf{r}_3 . \tag{10.7}$$

In terms of $g(r) = G(r) + 1$, this equation is

$$g(r_{12}) e^{\beta\phi(r_{12})} = 1 + n \int (1 - e^{\beta\phi(r_{13})}) g(r_{13}) [g(r_{32}) - 1]\, d\mathbf{r}_3 . \tag{10.8}$$

From Eq. (9.5), one finds that the PY approximation amounts to having $E(r)$ not equal to zero but as follows:

$$E(r) = -\left[\frac{S^2(r)}{2!} + \frac{S^3(r)}{3!} + \ldots\right]. \tag{10.9}$$

Nevertheless, one cannot say that the PY approximation is always better than the HNC approximation. In general, it yields good results for hard-sphere potentials.

In terms of graphs, both the HNC and PY approximations are correct to first-order diagram in density. To second order, the HYC approximation misses one graph and the PY approximation drops two.

Figure 6.5 illustrates the radial distribution function of argon at 84.4°K

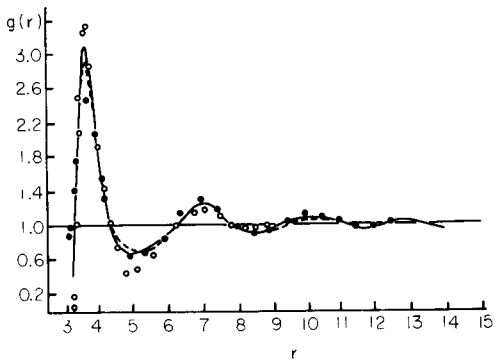


Fig. 6.5. Radial distribution function of argon. Black dots: neutron data at 84°K; open dots: X-ray data at 84.4°K, $n = 2.113 \times 10^{-2}$ Å^{-3}. (——) PY and (– – –) HNC approximations calculated by Khan (*22*).

determined by neutron scattering by Henshow (*20*) and by X-ray scattering by Eisenstein and Gingrich (*21*). The solid and dashed curves are evaluated by Khan based respectively on the PY and HNC approximations (*22*, *23*). The difference between the two approximations is small in this graph, but when thermodynamic functions are evaluated, the difference can be large (*22*, *23*). For example, for p/nkT, the experimental value is 0.004, the PY approximation gives 0.45, while the HNC approximation gives 1.12. In the evaluation, the potential function is chosen as follows:

$$\phi(r) = 4\varepsilon\left[\left(\frac{a}{r}\right)^{12} - \left(\frac{a}{r}\right)^{6}\right];$$

$$\varepsilon/k = 119.76 \quad °\text{K}, \qquad a = 3.405 \quad \text{Å}.$$

The theoretical curves depend very much on the choice of the potential functions. The temperature dependence also comes in. At $T = 149.3$°K, the PY approximation gives 0.465 for p/nkT and the HNC approximation gives 0.468, while the experimental value is 0.433.

6.11.* Born–Green Theory

The HNC and PY approximations are connected with what are known as the Born–Green and Kirkwood integral equations. For example, if we take the gradient of Eq. (10.3), we find

$$\frac{\partial}{\partial \mathbf{r}_1}[\log g(r_{12}) + \beta\phi(r)] = -n\beta\int \frac{\partial\phi(r_{13})}{\partial \mathbf{r}_1}\frac{g(r_{12})g(r_{13})g(r_{32})}{g(r_{12})}\,d\mathbf{r}_3 + \Delta(r),$$

where

$$\Delta(r) = n\int\left[\frac{\partial g(r_{13})}{\partial \mathbf{r}_1} - \frac{\partial}{\partial \mathbf{r}_1}\log g(r_{13}) - \beta\frac{\partial\phi(r_{13})}{\partial \mathbf{r}_1}\right][g(r_{32}) - 1]\,d\mathbf{r}_3$$

$$+ \beta n\int \frac{\partial\phi(r_{13})}{\partial \mathbf{r}_1} g(r_{13})g(r_{32})\,d\mathbf{r}_3 . \tag{11.1}$$

Integral equations similar to the above have been derived by Born and Green, Kirkwood, and others [cf. Eq. (11.4)]. The function $\Delta(r)$ represents the difference between the HNC and the Born–Green and Kirkwood equations. The PY equation (10.8) can also be brought into a similar form.

We shall now try to derive the Born–Green integral equation based on an analytical method instead of the graphical method used in the previous section. All these integral equations can be derived in several different ways, and their interrelationships are interesting. However, we shall omit the detailed discussions.

By differentiating Eq. (5.5), we find that the radial distribution function $g(r) = \rho_2(r)/n^2$ satisfies

$$-kT\frac{\partial}{\partial \mathbf{r}_2}\log g(\mathbf{r}_1, \mathbf{r}_2) = \frac{\int\cdots\int \frac{\partial\Phi}{\partial \mathbf{r}_2} e^{-\beta\Phi}\,d\mathbf{r}_3\cdots d\mathbf{r}_N}{\int\cdots\int e^{-\beta\Phi}\,d\mathbf{r}_3\cdots d\mathbf{r}_N} = {}^{12}\left\langle\frac{\partial\Phi}{\partial \mathbf{r}_2}\right\rangle. \tag{11.2}$$

The symbol ${}^{12}\langle\cdots\rangle$ on the right-hand side means the average taken at constant $\mathbf{r}_1$ and $\mathbf{r}_2$. Let us introduce the expression

$$g(\mathbf{r}_1, \mathbf{r}_2) = e^{-\beta w(r)}, \qquad r = |\mathbf{r}_2 - \mathbf{r}_1|. \tag{11.3}$$

$w(r)$ is called *the potential of average force*. As we see from Eq. (8.11), the first approximation for $w(r)$ is the potential $\phi(r)$.

Using $\Phi = \sum \phi_{ij}$, we evaluate the right-hand side of Eq. (11.2) as follows:

$$-kT\frac{\partial}{\partial \mathbf{r}_2}\log g(\mathbf{r}_1, \mathbf{r}_2)$$

$$= \frac{\partial\phi(\mathbf{r}_1, \mathbf{r}_2)}{\partial \mathbf{r}_2} + \int \frac{\rho_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)}{\rho_2(\mathbf{r}_1, \mathbf{r}_2)}\frac{\partial\phi(\mathbf{r}_2, \mathbf{r}_3)}{\partial \mathbf{r}_2}\,d\mathbf{r}_3, \tag{11.4}$$

which coincides with what we expect from Eq. (1.8) for equilibrium cases.

Using $f^{(1)} = \rho_l$ times Maxwell's distribution, and integrating Eq. (1.8) over the momenta, we find

$$\frac{\partial \rho_l}{\partial \mathbf{r}_i} + \sum_{j \neq i} \left[\frac{\rho_l}{kT} \frac{\partial \phi_{i,j}}{\partial \mathbf{r}_i} \right] + \frac{1}{kT} \int \frac{\rho_{l+1}}{kT} \frac{\partial \phi_{i,l+1}}{\partial \mathbf{r}_i} \, d\mathbf{r}_{i,l+1} = 0. \tag{11.5}$$

In order to have the integral equation in closed form, we assume

$$n^3 \rho_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = \rho_2(\mathbf{r}_1, \mathbf{r}_2) \rho_2(\mathbf{r}_2, \mathbf{r}_3) \rho_2(\mathbf{r}_3, \mathbf{r}_1), \tag{11.6}$$

which was first introduced by Kirkwood and called the *superposition approximation.*

Born and Green used this superposition approximation in Eq. (11.4) and derived the following integral equation:

$$\log g(r) = -\beta\phi(r) + \frac{\beta\pi n}{r} \int_0^\infty ds \times \int_{-s}^{s} \phi'(s)(s^2 - t^2)(t + r)[g(|r + t|) - 1]g(s) \, dt, \tag{11.7}$$

where $t = r_{13}$ and $s = r_{23}$. This is a nonlinear integral equation. Expanding $g(r)$ in a power series as in Eq. (8.11), it is possible to show that the equation is correct up to the third virial coefficient.

In order to treat the integral equation analytically, we assume the form

$$g(r) = e^{-\beta\phi(r) + h(r)}. \tag{11.8}$$

Linearizing the integral equation with respect to $h(r)$, we find

$$rh(r) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{\infty} \frac{\varepsilon^2 s[\beta(s)]^2 \sin rs}{\lambda - \varepsilon\beta(s)} \, ds, \tag{11.9}$$

where

$$\begin{aligned} \lambda^{-1} &= (2\pi)^{3/2} n, \\ \varepsilon - 1 &= \frac{\int_0^\infty r^2 h(r) f(r) \, dr}{\int_0^\infty r^2 f(r) \, dr}, \\ f(r) &= e^{-\beta\phi(r)} - 1, \\ s\beta(s) &= (2\pi)^{-1/2} \int_{-\infty}^{\infty} rf(r) \sin rs \, dr. \end{aligned} \tag{11.10}$$

It has been found that, in the case of hard-sphere molecules, Eq. (11.9) is correct to the third virial coefficient. As is given in Eq. (1.12) of Chapter 5, the second virial coefficient is $B = 4v_0$, with v_0 the molecular volume. The

third virial coefficient C is $0.4B^2$. The fourth virial coefficient D evaluated from Eq. (11.9) turned out to be $0.3958B^3$, whereas the correct value is $0.2869B^3$.

It is interesting to see the behavior of the solution to Eq. (11.9) at different temperatures. At high temperatures, $\beta(s)$ is small. Since λ^{-1} is proportional to the number density, the denominator does not vanish at low densities. As the temperature is lowered and the density is increased, there will be a temperature at which the following will hold:

$$\lambda = \varepsilon\beta(s). \tag{11.11}$$

Figure 6.6 illustrates $\beta(s)$ as a function of s. The straight line 2 borders the

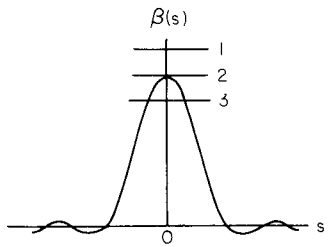


Fig. 6.6. Plot of $\beta(s)$ as a function of s. The straight lines 1, 2, and 3 correspond to different values of λ/ε.

region in which the above equation has real solutions: Below this line, there will be at least two real roots. Therefore, we expect a discontinuity in $\rho_2(r)$ starting at the line 2. This discontinuity may be connected to condensation.

For a discussion of condensation, it is convenient to consider s as a complex variable (see Section 5.7). Then, Cauchy's integral theorem indicates that the integral of Eq. (11.9) is determined by the residues at the roots of Eq. (11.11). When two real roots appear on the real axis, condensation takes place.

Problems

1. Consider an N-particle system with the potential

$$\Phi(r^N) = \xi \sum_{i=2}^{N} \phi(r_{ij}) + \sum_{i=2}^{N} \sum_{j>i}^{N} \phi(r_{ij}),$$

where ξ is a coupling parameter that varies from 0 to 1. By differentiating the equation that defines $g(r)$, show that

$$\frac{\partial}{\partial \xi} \log g(r_{12}, \xi) = -\beta\phi(r_{12}) g(r_{12}, \xi) - n\beta \int \left[\frac{g_3(r_{12}, r_{13}, r_{23})}{g(r_{12}, \xi)} - 1 \right] \phi(r_{13}) \, g(r_{12}, \xi) \, d\mathbf{r}_3,$$

where $g_3(r_{12}, r_{13}, r_{23})$ is the correlation function for the triplet distribution function. Here, r_{ij} is the relative distance between the molecules i and j.

2. Assuming the superposition approximation for the triplet correlation function and integrating with respect to ξ, derive from the above equation the following equation:

$$\log g(r_{12}, \xi) = -\xi\beta\phi(r_{12}) - n\beta \int_0^\xi \int \phi(r_{13}) g(r_{13}, \xi) [g(r_{23}) - 1] 4\pi r_{13}^2 \, dr_{13} \, d\xi.$$

This equation is due to Kirkwood and Boggs.

3. Show that the Born–Green equation may be written in terms of the coupling parameter as follows:

$$\frac{\partial}{\partial r_{12}} \log g(r_{12}, \xi) = -\beta\xi \frac{\partial\phi(r_{12})}{r_{12}} - n\beta\xi \int_0^\xi \int \frac{\partial\phi(r_{13})}{\partial r_{13}} g(r_{13}, \xi) g(r_{23}) \times 4\pi r_{13}^2 \, dr_{13}.$$

4. Compare the Kirkwood–Boggs equation (Problem 2) with the Born–Green equation by differentiating the former.

5. Show that the specific heat at constant volume is given by

$$C_v = \frac{3Nk}{2} + \frac{Nn}{2} \int_0^\infty \phi(r) \frac{\partial g(r)}{\partial T} 4\pi r^2 \, dr.$$

REFERENCES

1. J. G. Kirkwood, *J. Chem. Phys.* **14**, 180 (1946); **15**, 72 (1947).
2. R. W. Zwanzig, J. G. Kirkwood, K. Stripp, and I. Oppenheim, *J. Chem. Phys.* **21**, 2050 (1953).
3. J. G. Kirkwood and J. Ross, *J. Chem. Phys.* **22**, 1094 (1954).
4. N. N. Bogoliubov, *J. Phys. (USSR)* **10**, 265 (1946).
5. H. S. Green, "The Molecular Theory of Fluids." North-Holland, Amsterdam, 1952.
6. G. H. A. Cole, "An Introduction to the Statistical Theory of Classical Dense Fluids." Pergamon, London, 1967.
7. J. E. Mayer, *J. Chem. Phys.* **15**, 187 (1947).
8. J. G. Kirkwood, F. P. Buff, and M. S. Green, *J. Chem. Phys.* **17**, 988 (1949).
9. N. S. Gingrich, *Rev. Mod. Phys.* **15**, 90 (1943). For the radial distribution function near the triple point see: N. S. Gingrich and C. W. Thompson, *J. Chem. Phys.* **36**, 2398 (1962).
10. J. de Boer, *Rep. Progr. Phys.* **12**, 305 (1948).
11. G. S. Rushbrooke and H. I. Scoins, *Proc. Roy. Soc. (London), Ser. A* **216**, 203 (1953).
12. A. Isihara, *Phys. Rev.* **178**, 412 (1969).

12a. L. S. Ornstein and F. Zernike, *Proc. Nat. Acad Sci. U.S.A.* **17**, 793 (1914).

13. E. Meeron, *J. Math. Phys.* **1**, 192 (1960).
14. T. Morita, *Progr. Theoret. Phys.* **23**, 829 (1960); **41**, 239 (1969).

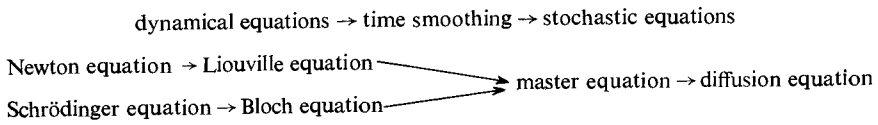
15. D. L. Bowers and E. E. Salpeter, *Phys. Rev.* **119**, 1180 (1960).
16. A. Isihara, *Phys. Rev.* **178**, 412 (1969).
17. J. M. J. van Leeuwen, J. Groeneveld, and J. de Boer, *Physica* **25**, 792 (1959).
18. J. K. Percus and G. J. Yevick, *Phys. Rev.* **110**, 1 (1958).
19. H. L. Frisch and J. L. Lebowitz (eds.) "The Equilibrium Theory of Classical Fluids." Benjamin, New York, 1964.
20. D. G. Henshow, *Phys. Rev.* **105**, 976 (1957).
21. A. Eisenstein and N. S. Gingrich, *Phys. Rev.* **62**, 261 (1942).
22. A. A. Khan, *Phys. Rev.* **134**, A367 (1964).
23. A. A. Broyles *et al.*, *J. Chem. Phys.* **37**, 2462 (1962).

7

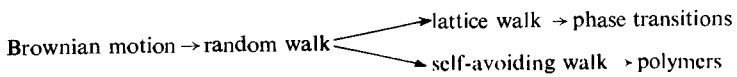
Brownian Motion

The theory of Brownian motion is important not only in statistical mechanics but also in many other branches of theoretical physics. As we shall discuss later, when the Hamiltonian does not include interactions, the Schrödinger equation and the Bloch equation have the same mathematical structure as the diffusion equation of a Brownian particle. When interactions are included, the solutions of these equations may be found by a perturbation method starting with the solution of the "diffusion equation." The Liouville and the Dirac equations may also be treated similarly. Feynman's diagram method is a typical example.

In irreversible statistical mechanics, it is a fundamental problem to derive irreversibility from reversible dynamics. In this respect, the theory of Brownian motion plays a unique role. In general, the process from dynamical equations to stochastic equations takes the following route:


dynamical equations → time smoothing → stochastic equations
Newton equation → Liouville equation
Schrödinger equation → Bloch equation
master equation → diffusion equation

Concerning equilibrium statistical mechanics, the theory of Brownian motion has other interesting and important connections with the following seemingly different subjects:



The fascinating development of the theory of Brownian motion is briefly discussed in the next section. We shall then study the Markov method for random walk processes. After the treatment of random walks on lattices, Einstein's viscosity coefficient and Stokes's friction coefficient will be derived. We then adopt a phenomenological description based on the classical Langevin equation to find some important stochastic relations. This description will be followed by a molecular theory and a discussion of the Fokker–Planck equation. Our final subject in this chapter is the self-avoiding walk problem—a challenging problem that is still open.

In short, we shall study Brownian motion in the context indicated in the above two diagrams.

7.1. Random Walks and Brownian Motion

In 1828, Robert Brown published a paper concerning the rapid, irregular motion of pollens suspended in water. Later, he made additional, similar observations. However, he was not the person who discovered the phenomenon of Brownian motion. Before Brown, the phenomenon had been discovered by several predecessors. In particular, according to Brown himself, Bywater reached the conclusion around 1819 that, "Not only organic tissues, but even inorganic substances, consist of animated or irritable particles."

The discovery of Brownian motion led many scientists to believe that the particles were alive. Although Brown did not conclude that the particles were alive, he considered that the origin was in the "active molecules" themselves and not in the surrounding fluid. Brown did rediscover the motion in the sense that he clearly recognized the phenomenon in organic as well as in inorganic matter. This discovery aroused widespread interest.

In 1863, E. C. Wiener interpreted the Brownian motion in terms of a molecular kinetic theory. Shortly after this, Gouy confirmed the molecular kinetic viewpoint by careful experiments.

The year 1905 was made important for physics by Einstein's presentation of his theory of Brownian motion and his theory of relativity (*1*). Einstein's contribution to the former is twofold: One is the derivation of the diffusion equation, and the other is what we now call Einstein's relation, which relates the diffusion constant to temperature. Later, theories with more dynamical character were developed by Smoluchowski, Langevin, Ornstein, Uhlenbeck, and others.

The theory of Brownian motion is closely related to that of *random walks*. Instead of starting with a diffusion equation, we shall therefore first

consider a random walk process and look for the probability that a (Brownian) particle is at $\mathbf{r}$ after N displacements $\mathbf{r}_i$:

$$\mathbf{r} = \sum_{i=1}^{N} \mathbf{r}_i. \tag{1.1}$$

Let us introduce the probability $\tau_i(\mathbf{r}_i)\, d\mathbf{r}_i$ that the ith displacement lies between $\mathbf{r}_i$ and $\mathbf{r}_i + d\mathbf{r}_i$. If each step is independent of the others, the probability that a Brownian particle is at $\mathbf{r}$ after N displacements is given by

$$W_N(\mathbf{r}) = \int \cdots \int \delta\Big(\sum \mathbf{r}_i - \mathbf{r}\Big) \prod_{i=1}^{N} \tau_i(\mathbf{r}_i)\, d\mathbf{r}_i, \tag{1.2}$$

where δ is the Dirac delta function. Thus, introducing

$$A_N(\mathbf{q}) = \prod_{i=1}^{N} \int_{-\infty}^{\infty} \tau_j(\mathbf{r}_j) \exp(i\mathbf{q} \cdot \mathbf{r}_j)\, d\mathbf{r}_j, \tag{1.3}$$

we find $W_N(\mathbf{r})$ in the form

$$W_N(\mathbf{r}) = \frac{1}{8\pi^3} \int_{-\infty}^{\infty} \exp(-i\mathbf{q} \cdot \mathbf{r}) A_N(\mathbf{q})\, d\mathbf{q}. \tag{1.4}$$

If each displacement is arbitrary in its direction, we can introduce

$$\tau_j = \frac{1}{2\pi b_j} \delta({r_j}^2 - {b_j}^2). \tag{1.5}$$

Then,

$$\begin{aligned} A_N(\mathbf{q}) &= \prod_{j=1}^{N} \frac{1}{2\pi b_j} \int_{-\infty}^{\infty} \exp(i\mathbf{q} \cdot \mathbf{r}_j)\, \delta({r_j}^2 - {b_j}^2)\, d\mathbf{r}_j \\ &= \prod_{j=1}^{N} \frac{\sin qb_j}{qb_j}. \end{aligned} \tag{1.6}$$

Thus,

$$W_N(\mathbf{r}) = \frac{1}{2\pi^2 r} \int_0^{\infty} q\, dq \sin qr \prod \left(\frac{\sin qb_j}{qb_j}\right). \tag{1.7}$$

In particular, if all the b_j are equal to b and if $N \gg 1$, we may use

$$\begin{aligned} A_N(\mathbf{q}) = \lim_{N\to\infty} \left(\frac{\sin qb}{qb}\right)^N &= \lim \left(1 - \frac{1}{6} q^2 b^2 + \cdots\right)^N \\ &= \exp\left(-\frac{Nq^2 b^2}{6}\right). \end{aligned} \tag{1.8}$$

Introducing Eq. (1.8) into Eq. (1.7), we obtain $W_N(\mathbf{r}) \to W(\mathbf{r})$:

$$W(\mathbf{r}) = \frac{1}{(\frac{2}{3}\pi N b^2)^{3/2}} \exp\left(-\frac{3r^2}{2Nb^2}\right). \tag{1.9}$$

The above method is due to Markov. Note that it is applicable to any other vector besides displacement that has the form of Eq. (1.1). Note also that

$$\langle r^2 \rangle = Nb^2, \tag{1.10}$$

a result which plays an important role in the theory of rubber elasticity.

In the limit $N \to \infty$, one can introduce a continuous time variable t by $N = nt$, n being the time rate of steps. Defining

$$D = \tfrac{1}{6} n \langle r^2 \rangle, \tag{1.11}$$

we can rewrite Eq. (1.9) as follows:

$$W(\mathbf{r}) = \frac{1}{(4\pi Dt)^{3/2}} \exp\left(-\frac{r^2}{4Dt}\right). \tag{1.12}$$

Let Δt be an interval of time long enough for a Brownian particle to undergo a large number of displacements, but short enough for the total mean square increment $\langle (\Delta r)^2 \rangle$ in the period to remain small. Then, the probability distribution function $W(\mathbf{r}, t + \Delta t)$ at $t + \Delta t$ is related to that at t by

$$W(\mathbf{r}, t + \Delta t) = \int_{-\infty}^{\infty} W(\mathbf{r} - \Delta\mathbf{r}, t) K(\Delta\mathbf{r}, \Delta t)\, d(\Delta\mathbf{r}). \tag{1.13}$$

Expanding both sides in Taylor series in $\Delta\mathbf{r}$, we arrive at the diffusion equation:

$$\frac{\partial W}{\partial t} = D \nabla^2 W. \tag{1.14}$$

The solution of this diffusion equation corresponding to the initial condition that there was a source at $t = 0$ is simply given by Eq. (1.12).

We note also that Eq. (1.12) is a result of the Fourier transformation of Eq. (1.8). Rewriting Eq. (1.8), we find

$$\tilde{W}(\mathbf{q}) = \exp(-q^2 Dt). \tag{1.15}$$

The Fourier transform of Eq. (1.15) in space–time is important. Replacing t by $|t|$ in Eq. (1.15), we find

$$\begin{aligned} S(q, \omega) &= \int_{-\infty}^{\infty} \tilde{W}(\mathbf{q}) e^{-i\omega t}\, dt \\ &= \frac{1}{\pi} \frac{Dq^2}{\omega^2 + (Dq^2)^2}. \end{aligned} \tag{1.16}$$

The function $S(q, \omega)$ plotted against ω is Lorentzian with a half-width of Dq^2. Thus, if $S(q, \omega)$ is obtained for a number of q^2 values, one can determine the diffusion constant D. Its accurate determination can be made in the small (q, ω) range, but neutron scattering experiments for liquids usually correspond to values of (q, ω) too large to make an accurate determination of D. Light scattering measurements are not suitable for the determination since the scattering is coherent.

In this respect, we recall that Eq. (1.12) is the result that is valid for large $(\mathbf{r}, t)$ values. For small $(\mathbf{r}, t)$, we must look into the details of the diffusion process. In liquids, a particle may be considered to stay in each position for awhile before making the next diffusion step. The time for each "stop" relative to that of "flight" is characteristic of the liquid state.

7.2.* Random Walks on Lattices

Let us find the probability $P_N(\mathbf{R})$ that a random walker reaches a lattice position $\mathbf{R}$,

$$\mathbf{R} = N_1\mathbf{a}_1 + N_2\mathbf{a}_2 + N_3\mathbf{a}_3,$$

after $N = N_1 + N_2 + N_3$ steps on a lattice with unit vectors $\mathbf{a}_i$. We assume that the lattice is periodic with a periodicity L, so that no boundary condition comes in. The probability satisfies the recurrence formula

$$P_{N+1}(\mathbf{R}) = \sum_{\mathbf{R}'} W(\mathbf{R} - \mathbf{R}')P_N(\mathbf{R}'), \tag{2.1}$$

where $W(\mathbf{R} - \mathbf{R}')$ is the probability of transition from $\mathbf{R}'$ to $\mathbf{R}$. We introduce a generating function (*1a*)

$$\Phi(\mathbf{R}, z) = \sum_{N=0}^{\infty} z^N P_N(\mathbf{R}), \tag{2.2}$$

and note that

$$\Phi(\mathbf{R}, z) - z \sum_{\mathbf{R}'} W(\mathbf{R} - \mathbf{R}')\Phi(\mathbf{R}', z) = \delta_{\mathbf{R},0}, \tag{2.3}$$

which expresses the fact that the walker starts walking from the origin:

$$P_0(\mathbf{R}) = \delta_{\mathbf{R},0} = \begin{cases} 1, & \mathbf{R} = 0, \\ 0, & \text{otherwise.} \end{cases} \tag{2.4}$$

Here, δ is the Kronecker delta.

Equation (2.3) may be solved by a Fourier transform method. Multiplying both sides by $\exp(2\pi i\mathbf{k} \cdot \mathbf{R}/L)$ and summing over $\mathbf{R}$, we find

$$\tilde{\Phi}(\mathbf{K}, z) - z\lambda(\mathbf{K})\tilde{\Phi}(\mathbf{K}, z) = 1, \tag{2.5}$$

with

$$\tilde{\Phi}(\mathbf{K}, z) = \sum_{\mathbf{R}} \exp\left[\frac{2\pi i\mathbf{K} \cdot \mathbf{R}}{L}\right]\Phi(\mathbf{R}, z); \tag{2.6}$$

$$\lambda(\mathbf{K}) = \sum_{\mathbf{R}} W(\mathbf{R}) \exp\left[\frac{(2\pi i\mathbf{K} \cdot \mathbf{R})}{L}\right]. \tag{2.7}$$

The Fourier inverse of $\tilde{\Phi}(\mathbf{K}, z)$ is

$$\Phi(\mathbf{R}, z) = \frac{1}{L^3} \sum_{\mathbf{K}} \frac{\exp(-2\pi i\mathbf{K} \cdot \mathbf{R}/L)}{1 - z\lambda(\mathbf{K})}, \tag{2.8}$$

which, in the limit of $L \to \infty$, gives

$$\lim_{L\to\infty} \Phi(\mathbf{R}, z) = \frac{1}{(2\pi)^3} \iiint_{-\pi}^{\pi} \frac{\exp(-i\boldsymbol{\theta} \cdot \mathbf{R})}{1 - z\lambda(\boldsymbol{\theta})}\, d\boldsymbol{\theta}. \tag{2.9}$$

The probability $P_N(\mathbf{R})$ is given by

$$P_N(\mathbf{R}) = \frac{1}{(2\pi)^3} \iiint_{-\pi}^{\pi} [\lambda(\boldsymbol{\theta})]^N \exp(-i\boldsymbol{\theta} \cdot \mathbf{R})\, d\boldsymbol{\theta}. \tag{2.10}$$

The transition probability $W(\mathbf{R})$ for a unit displacement is determined by the lattice structure. For a simple cubic lattice, we have

$$W(\mathbf{R}) = \begin{cases} \frac{1}{6}, & \text{for } \mathbf{R} = (0, 0, \pm 1), (0, \pm 1, 0), (\pm 1, 0, 0); \\ 0, & \text{otherwise.} \end{cases} \tag{2.11}$$

Therefore, we find

$$\begin{aligned} \lambda(\boldsymbol{\theta}) &= \tfrac{1}{6}(e^{i\theta_1} + e^{-i\theta_1} + e^{i\theta_2} + e^{-i\theta_2} + e^{i\theta_3} + e^{-i\theta_3}) \\ &= \tfrac{1}{3}(\cos\theta_1 + \cos\theta_2 + \cos\theta_3). \end{aligned} \tag{2.12}$$

Thus,

$$P_N(\mathbf{R}) = \frac{1}{(2\pi)^3} \iiint_{-\pi}^{\pi} \left[\frac{1}{3}(\cos\theta_1 + \cos\theta_2 + \cos\theta_3)\right]^N \exp(-i\boldsymbol{\theta} \cdot \mathbf{R})\, d\boldsymbol{\theta}, \tag{2.13}$$

and

$$\Phi(\mathbf{R}, z) = \frac{1}{(2\pi)^3} \iiint_{\pi}^{\pi} \frac{\exp(-i\boldsymbol{\theta} \cdot \mathbf{R})}{1 - \frac{1}{3}z(\cos\theta_1 + \cos\theta_2 + \cos\theta_3)}\, d\boldsymbol{\theta}. \tag{2.14}$$

In particular, for $R = 0$,

$$\Phi(0, z) = \frac{1}{(2\pi)^3} \iiint_{-\pi}^{\pi} \frac{d\theta_1\, d\theta_2\, d\theta_3}{1 - \frac{1}{3}z(\cos\theta_1 + \cos\theta_2 + \cos\theta_3)}. \tag{2.15}$$

$$\Phi(0, 1) \sim 1.5164. \tag{2.16}$$

On the other hand, for $R \to \infty$, the main contributions to the θ-integrals come from the neighborhood of the origin. Thus, for $R \to \infty$, we expand $\cos\theta$ and find

$$\Phi(R, z) \sim \frac{1}{(2\pi)^3} \iiint_{-\infty}^{\infty} \frac{\exp(-i\boldsymbol{\theta}\cdot\mathbf{R})\, d\boldsymbol{\theta}}{1 - z + \frac{1}{6}z(\theta_1^2 + \theta_2^2 + \theta_3^2)} \tag{2.17}$$

$$= \frac{3}{2\pi z R} \exp\left\{-3^{1/2} R \left[\frac{2(1-z)}{z}\right]^{1/2}\right\}, \tag{2.18}$$

where $R = |\mathbf{R}|$. Also, from Eq. (2.13), we obtain

$$P_N(\mathbf{R}) \to \left(\frac{3}{2\pi N}\right)^{3/2} \exp\left(-\frac{3}{2}\frac{R^2}{N}\right). \tag{2.19}$$

7.3.* Stokes Friction and Einstein Viscosity

When colloidal particles are dissolved in solutions, the viscosity increases with the concentration. At infinite dilution, the increase in the viscosity coefficient over that of the pure solvent may be considered to be characteristic of the colloidal particles. It is customary to express the increase by a quantity called *intrinsic viscosity* by

$$\lim_{c\to 0} \frac{\eta - \eta_0}{\eta_0 c} = [\eta], \tag{3.1}$$

where η is the viscosity of solution, η_0 is that for the pure solvent, and c is the weight concentration.

In 1906, Einstein published an important paper to relate $[\eta]$ to the dimension of a colloidal particle considered to be spherical. To evaluate the intrinsic viscosity, we consider a sphere of radius a at the origin in a laminar flow $\mathbf{v}^0$ which may be decomposed into two parts:

$$\mathbf{v}^0 = \mathbf{v}_1 + \mathbf{v}';$$

$$\mathbf{v}_1 = \begin{cases} \frac{-qy}{2}, \\ \frac{qx}{2}, \\ 0, \end{cases} \qquad \mathbf{v}' = \begin{cases} \frac{qy}{2}, \\ \frac{qx}{2}, \\ 0, \end{cases} \tag{3.2}$$

where $\mathbf{v}_1$ represents a rotation of the fluid and $\mathbf{v}'$ is a shear flow which causes viscosity. The fluid flow at the surface of the sphere is considered to be zero. The constant q represents the velocity gradient.

A spherical particle at the origin will rotate with angular velocity $q/2$ following the fluid motion. However, this motion is not enough to have an equivelocity distribution, and a local disturbance of the flow appears. The effect is the same if a force is put at the origin and causes the perturbation flow.

According to Oseen, a force $\mathbf{F}$ at the origin causes a perturbation $\mathbf{v}$ of flow. If $F_y = F_z = 0$, the perturbation at a point $\mathbf{r}$ is given by

$$\mathbf{v} = \begin{cases} v_x = \frac{F_x}{8\pi\eta_0}\left(\frac{1}{r} + \frac{x^2}{r^3}\right), \\ v_y = \frac{F_x}{8\pi\eta_0}\frac{xy}{r^3}, \\ v_z = \frac{F_x}{8\pi\eta_0}\frac{xz}{r^3}. \end{cases} \tag{3.3}$$

The *Oseen formula* (3.3) yields the average of the perturbation flow on the surface of a sphere with its center at the origin as follows:

$$\langle v_x\rangle = \frac{F_x}{8\pi\eta_0}\left(\frac{1}{a} + \frac{1}{a}\langle\cos^2\theta\rangle\right)$$

$$= \frac{F_x}{6\pi\eta_0 a}, \tag{3.4}$$

where a is the radius of the sphere.

Equation (3.4) gives the Stokes force

$$F_x = 6\pi\eta_0 a\langle v\rangle \tag{3.5}$$

acting on the fluid. Therefore, the sphere experiences a resistance $-6\pi\eta_0 a\langle v\rangle$ if it moves with a velocity $\langle\mathbf{v}\rangle$. This result was first found by Stokes in 1850.

Let us now derive the Einstein viscosity formula for spherical particles (*2*). To reproduce the perturbation of flow due to a sphere at the origin, it is necessary to consider a doublet force at the origin instead of a single force.

Thus, we introduce a positive force at $(x = 0, y = -\varepsilon, z = 0)$ and a negative force at $(x = 0, y = \varepsilon, z = 0)$. The former produces the flow

$$\mathbf{v}_{\mathrm{I}} = \begin{cases} v_{\mathrm{I}x} = A\left(-\dfrac{y}{r^3} - \dfrac{3x^2y}{r^5}\right), \\ v_{\mathrm{I}y} = A\left(\dfrac{x}{r^3} - \dfrac{3xy^2}{r^5}\right), \\ v_{\mathrm{I}z} = -A\dfrac{3xyz}{r^5}, \end{cases} \tag{3.6a}$$

while the latter gives

$$\mathbf{v}_{\mathrm{II}} = \begin{cases} v_{\mathrm{II}x} = A\left(\dfrac{y}{r^3} - \dfrac{3x^2y}{r^5}\right), \\ v_{\mathrm{II}y} = A\left(-\dfrac{x}{r^3} - \dfrac{3xy^2}{r^5}\right), \\ v_{\mathrm{II}z} = -A\dfrac{3xyz}{r^5}. \end{cases} \tag{3.6b}$$

Here, A is an undetermined constant related to the strength of the doublet. To these two velocity fields we add a third one which is derived from the potential

$$\Phi = B\frac{\partial^2}{\partial x\,\partial y}\left(\frac{1}{r}\right). \tag{3.7}$$

This yields

$$\mathbf{v}_{\mathrm{III}} = \begin{cases} v_{\mathrm{III}x} = B\left(\dfrac{3y}{r^5} - \dfrac{15}{r^7}x^2y\right), \\ v_{\mathrm{III}y} = B\left(\dfrac{3x}{r^5} - \dfrac{15}{r^7}xy^2\right), \\ v_{\mathrm{III}z} = -B\dfrac{15xyz}{r^7}. \end{cases} \tag{3.8}$$

When all these fields are superposed, we find that the choice

$$B = -\tfrac{1}{2}Aa^2, \tag{3.9}$$

with a for the radius of the sphere, yields a perturbation

$$\mathbf{v}' = \begin{cases} v_x' = -\dfrac{6}{5}\dfrac{Ay}{a^3}, \\ v_y' = -\dfrac{6}{5}\dfrac{Ax}{a^3}, \\ v_z' = 0, \end{cases} \tag{3.10}$$

on the surface of the sphere. Comparing this with Eq. (3.2), we discover that, if A is chosen such that

$$-\frac{6}{5}\frac{A}{a^3} = \frac{1}{2}q, \tag{3.11}$$

then the combined perturbation $\mathbf{v}'$ reduces to that of a sphere. Equation (3.11) gives

$$A = -\frac{5}{12}qa^3. \tag{3.12}$$

The velocity $\mathbf{v}'$ has been superposed so that the result correctly represents the original fluid flow on the surface of the sphere. At large distances from the sphere, the effects of the sphere should vanish.

We observe that the third flow $\mathbf{v}_{\mathrm{III}}$ is small at large distances. Thus, combining the fields $\mathbf{v}_{\mathrm{I}}$ and $\mathbf{v}_{\mathrm{II}}$, we find the perturbation $\mathbf{v}'$ to be

$$\mathbf{v}' = \begin{cases} v_x' = \left(\dfrac{5}{2}qa^3\right)\dfrac{x^2y}{r^5}, \\ v_y' = \left(\dfrac{5}{2}qa^3\right)\dfrac{xy^2}{r^5}, \\ v_z' = \left(\dfrac{5}{2}qa^3\right)\dfrac{xyz}{r^5}. \end{cases} \tag{3.13}$$

The original solvent flow is caused by a force $\eta_0 q$ per unit surface acting oppositely in two parallel planes perpendicular to the x axis. Here η_0 is the viscosity of the pure solvent. The energy dissipation per unit time per unit volume in connection with this motion is $\eta_0 q^2$. On the other hand, the energy dissipation due to small spheres with number density n may be obtained by superposing those from individual spheres as follows:

$$2\iint_{x=-h}^{x=h}\left(\frac{5}{2}qa^3\right)n\frac{xy^2}{r^5}(\eta_0 q)\,dx\,dy = \left(\frac{5}{2}n\eta_0 q^2\right)\left(\frac{4\pi a^3}{3}\right). \tag{3.14}$$

$x = h$ and $x = -h$ represent the locations of the parallel planes.

Equating the right-hand side of Eq. (3.14) with $\eta_0 q^2$ we arrive at a specific viscosity increase

$$\frac{\eta - \eta_0}{\eta_0} = \frac{5n}{2}\left(\frac{4\pi a^3}{3}\right), \tag{3.15}$$

which is called the *Einstein viscosity formula.*

7.4. Langevin's Equation

A Brownian particle may be considered to be subject to a fluctuating force exerted by the surroundings in addition to a stochastic resistance proportional to the velocity of the particle. The *Langevin equation* has been introduced to describe such a motion for this situation:

$$\dot{\mathbf{v}} = -\gamma \mathbf{v} + \mathbf{X}(t), \tag{4.1}$$

where γ is a friction coefficient per unit mass. For a spherical particle of radius, a, γ is given by Stokes's law as follows:

$$\gamma = \frac{6\pi\eta a}{m} = \frac{\zeta}{m}, \tag{4.2}$$

where η is the viscosity of the liquid in which the Brownian particle moves and m is the mass. The term $m\mathbf{X}(t)$ is a fluctuating force, considered to be independent of the velocity $\mathbf{v}$, and satisfies the condition that its time average is zero:

$$\langle \mathbf{X}(t) \rangle = 0. \tag{4.3}$$

The formal solution of the Langevin equation is

$$\mathbf{v}(t) - \mathbf{v}(0)e^{-\gamma t} = e^{-\gamma t}\int_0^t e^{\gamma x}\mathbf{X}(x)\,dx. \tag{4.4}$$

The integral on the right-hand side may be rewritten as follows:

$$\sum_j e^{\gamma j\,\Delta t}\int_{j\,\Delta t}^{(j+1)\,\Delta t} \mathbf{X}(x)\,dx = \sum_j e^{\gamma j\,\Delta t}\overline{\mathbf{X}}\,\Delta t.$$

Thus, following the method discussed in Section 7.1, one can find the distribution of $\mathbf{v}(t)$ about $\mathbf{v}(0)e^{-\gamma t}$. This velocity distribution may be assumed to approach the Maxwell distribution for $\gamma t \gg 1$. This requirement leads us to the conclusion that the probability distribution for $\overline{\mathbf{X}}\,\Delta t$ is Gaussian

$$\tau(\overline{\mathbf{X}}\,\Delta t) = (4\pi q\,\Delta t)^{-3/2}\exp\left(\frac{-|\overline{\mathbf{X}}\Delta t|^2}{4q\,\Delta t}\right), \tag{4.5}$$

where $q = \gamma kT/m$.

Similarly, one can find the distribution for the displacement. Integrating Eq. (4.4), one obtains

$$\mathbf{r}(t) - \gamma^{-1}\mathbf{v}(0)(1 - e^{-\gamma t}) = \int_0^t dy\, e^{-\gamma y}\int_0^y dx\, e^{\gamma x}\mathbf{X}(x), \tag{4.6}$$

where the condition $\mathbf{r}(0) = 0$ is used. By integrating Eq. (4.6) by parts, we find

$$\mathbf{r}(t) - \gamma^{-1}\mathbf{v}(0)(1 - e^{-\gamma t}) = -\gamma^{-1}e^{-\gamma t}\int_0^t e^{\gamma x}\mathbf{X}(x)\,dx + \gamma^{-1}\int_0^t \mathbf{X}(x)\,dx. \tag{4.7}$$

For convenience, let us define a function $\psi(x)$ by

$$\psi(x) = \gamma^{-1}(1 - e^{\gamma(x-t)}). \tag{4.8}$$

Then, Eq. (4.7) becomes

$$\mathbf{r}(t) - \mathbf{v}(0)\psi(0) = \int_0^t \psi(x)\mathbf{X}(x)\,dx. \tag{4.9}$$

The time interval $(0, t)$ may be divided into a large number of subintervals.

We notice that the term on the left-hand side is a result of a random walk process:

$$\sum_j \psi(j\,\Delta t)\int_{j\Delta t}^{(j+1)\Delta t} \mathbf{X}(x)\,dx = \sum_j \psi(j\,\Delta t)\overline{\mathbf{X}}\,\Delta t.$$

From Eq. (4.5), the distribution of $\overline{\mathbf{X}}\,\Delta t$ is Gaussian if the system approaches equilibrium for $t \to \infty$. Furthermore,

$$\int_0^t \psi^2(x)\,dx = \frac{1}{\gamma^2}\int_0^t (1 - e^{\gamma(x-t)})^2\,dx = \frac{1}{2\gamma^3}(2\gamma t - 3 + 4e^{-\gamma t} - e^{-2\gamma t}). \tag{4.10}$$

Therefore, for a Brownian particle that started moving from the origin $\mathbf{r} = 0$ with an initial velocity $\mathbf{v}(0)$, the probability distribution $W(\mathbf{r}, t)$ is given, from Eqs. (1.4) and (4.5), by

$$W(\mathbf{r}, t) = \frac{1}{[4\pi q\int_0^t \psi^2(x)\,dx]^{3/2}}\exp\frac{-[\mathbf{r} - \mathbf{v}(0)\psi(0)]^2}{4q\int_0^t \psi^2(x)\,dx}, \tag{4.11}$$

where the quantity

$$D = \lim_{t\to\infty} \frac{q}{t}\int_0^t \psi^2(x)\,dx = \frac{kT}{m\gamma} = \frac{kT}{\zeta} \tag{4.12}$$

may be interpreted as the "diffusion constant" for the process. From Eq. (4.1), $1/m\gamma$ is the terminal velocity per unit force and is called *mobility*. Thus, we obtain

$$D = kT \times \text{mobility},$$

which is known as the *Einstein relation*. This definition is in agreement with Eq. (1.11) because, for $t \to \infty$,

$$\int_0^\infty \psi^2(x)\,dx = \frac{t}{\gamma^2}. \tag{4.13}$$

Also, from

$$\langle \mathbf{r}^2 \rangle = \frac{\langle v^2(0)\rangle}{\gamma^2}(1 - e^{-\gamma t})^2 + \frac{3kT}{m\gamma^2}(2\gamma t - 3 + 4e^{-\gamma t} - e^{-2\gamma t}), \tag{4.14}$$

we find, for $t \to \infty$,

$$\langle \mathbf{r}^2 \rangle = \frac{6kT}{m\gamma}\, t. \tag{4.15}$$

The proportionality to t of $\langle \mathbf{r}^2 \rangle$ is of the same nature as the random walk result of Eq. (1.11). Thus, $W(\mathbf{r}, t)$ reduces to $W(\mathbf{r})$ of Eq. (1.12) with the identification of Eq. (4.12).

On the other hand, for short times, we find

$$\langle \mathbf{r}^2 \rangle = \langle \mathbf{v}^2(0)\rangle t^2, \tag{4.16}$$

which is simply what we expect from dynamics.

We have derived the probability distribution of finding a Brownian particle at $\mathbf{r}$. It is clear that we could also obtain a distribution law for the velocity because Eq. (4.4) is of the structure of Eq. (4.9).

The Langevin equation in the presence of an external force is

$$\frac{d\mathbf{v}}{dt} + \gamma\mathbf{v} = \mathbf{X}(t) + \mathbf{F}(t). \tag{4.17}$$

If we define

$$\langle \mathbf{r}(t)\rangle = \int_{-\infty}^{\infty} \chi(t, t')\mathbf{F}(t')\,dt', \tag{4.18}$$

then $\chi(t, t')$ satisfies the Green's-function equation

$$\frac{d^2\chi}{dt^2} + \gamma\dot{\chi} = \delta(t - t'). \tag{4.19}$$

In particular, if χ depends only on $t - t'$ and has a Fourier transform

$$\tilde{\chi}(\omega) = \int_{-\infty}^{\infty} \chi(t - t')e^{i\omega(t-t')}\,d(t - t'), \tag{4.20}$$

$\tilde{\chi}(\omega)$ is found to be

$$\tilde{\chi}(\omega) = \frac{-1}{\omega^2 + \gamma i\omega}, \qquad \omega \neq 0. \tag{4.21}$$

$\tilde{\chi}(\omega)$ is called the *response function*. Using Eq. (4.21) in Eq. (4.20), we find

$$\chi(t - t') = \frac{(-1)}{2\pi}\int_{-\infty}^{\infty} d\omega\, \frac{e^{-i\omega(t-t')}}{\omega^2 + \gamma i\omega}, \tag{4.22}$$

where we perform a contour integral in the upper half of complex ω-plane.

The response function $\tilde{\chi}(\omega)$ may be complex. The real and the imaginary parts, $\tilde{\chi}'$ and $\tilde{\chi}''$, are related to each other by the *Kramers–Kronig relations*:

$$\begin{aligned} \tilde{\chi}'(\omega) &= P\int_{-\infty}^{\infty} \frac{d\omega'}{\pi}\frac{\tilde{\chi}''(\omega')}{\omega' - \omega}, \\ \tilde{\chi}''(\omega) &= -P\int_{-\infty}^{\infty} \frac{d\omega'}{\pi}\frac{\tilde{\chi}'(\omega')}{\omega' - \omega}, \end{aligned} \tag{4.23}$$

where P is the principal value. These relations may be proved by first expressing $\tilde{\chi}(\omega)$ by a Cauchy integral of itself and then making use of the identity

$$\lim_{\varepsilon\to 0} \frac{1}{x \mp i\varepsilon} = P\frac{1}{x} \pm \pi i\delta(x). \tag{4.24}$$

7.5. Friction Coefficient of a Brownian Particle

The description of Brownian motion by the Langevin equation is simple and instructive. However, since it is phenomenological, many efforts have been made to develop a molecular theory (*3–5*). This is not a simple problem, and we present here only a somewhat simplified version of the theoretical treatments.

Let us consider a system of N identical "field" particles plus a Brownian particle. The total distribution function will be denoted by $f^{(N+1)}(t) = f^{(N+1)}(\{N\}, (1); t)$, where the symbol $\{N\}$ represents the coordinates $\mathbf{r}_1, \ldots, \mathbf{r}_N$ and the momenta $\mathbf{p}_1, \ldots, \mathbf{p}_N$ of the N field particles, and $(1) = (\mathbf{R}, \mathbf{P})$ is for the Brownian particle. This distribution function is normalized such that

$$\int f^{(N+1)}\, d\{N\} = N!. \tag{5.1}$$

Since the Brownian particle is considered to be distinguishable from the field particles, let us introduce two reduced distribution functions:

$$\begin{aligned} f^{(1+1)}(\{1\}, (1); t) &= \frac{1}{(N-1)!} \int f^{(N+1)}\, d\{N-1\}, \\ f_B((1)) &= \frac{1}{N} \int f^{(1+1)}\, d\{1\}. \end{aligned} \tag{5.2}$$

Here, $f^{(1+1)}$ is the pair distribution function of a field particle and the Brownian particle, and f_B is the distribution function of the Brownian particle.

The total distribution function satisfies the Liouville equation:

$$\frac{\partial f^{(N+1)}}{\partial t} - i(L + L_B) f^{(N+1)} = 0, \tag{5.3}$$

where

$$\begin{aligned} -iL &= \sum_{j=1}^{N} \left(\frac{\mathbf{p}_j}{m} \cdot \frac{\partial}{\partial \mathbf{r}_j} + m\mathbf{X}_j \cdot \frac{\partial}{\partial \mathbf{p}_j} \right), \\ -iL_B &= \frac{\mathbf{P}}{M} \cdot \frac{\partial}{\partial \mathbf{R}} + M\mathbf{X} \cdot \frac{\partial}{\partial \mathbf{P}}, \end{aligned} \tag{5.4}$$

are called the Liouville operators. The imaginary unit i is introduced into these definitions in order to make the operators be self-adjoint. The forces $m\mathbf{X}_i$ and $M\mathbf{X}$ on the ith field particle and the Brownian particle are given by the interaction potentials as follows:

$$\begin{aligned} m\mathbf{X}_i &= -\sum_{j \neq i} \frac{\partial \phi_{ij}}{\partial \mathbf{r}_i} - \frac{\partial \Phi(\mathbf{r}_i - \mathbf{R})}{\partial \mathbf{r}_i}, \\ M\mathbf{X} &= -\frac{\partial}{\partial \mathbf{R}} \cdot \sum_i \Phi(\mathbf{r}_i - \mathbf{R}). \end{aligned} \tag{5.5}$$

In terms of these operators, the formal solution of the Liouville equation is given as follows (see Section 13.1 for details):

$$f^{(N+1)}(t) = \int_0^t e^{i(t-s)L}(iL_B) f^{(N+1)}(s)\, ds, \tag{5.6}$$

as can be confirmed by differentiation.

Let us now try to derive an irreversible equation for the Brownian particle. The differential equation to be satisfied by $f_B(t)$ is obtained by integrating the Liouville equation (5.3) over all the field particles. The result is one of the BBGKY hierarchy of equations given by Eq. (1.8) of Chapter 6 for $l = 1$. In our present notation, it is

$$\frac{\partial f_B}{\partial t} + \frac{\mathbf{P}}{M} \cdot \frac{\partial f_B}{\partial \mathbf{R}} = \int \frac{\partial \Phi(\mathbf{r} - \mathbf{R})}{\partial \mathbf{R}} \cdot \frac{\partial f^{(1+1)}}{\partial \mathbf{P}}\, d\{1\}. \tag{5.7}$$

As we have discussed before, this reduced Liouville equation does not represent irreversibility. It is still a dynamical equation. Irreversibility may be introduced only through further theoretical assumptions.

Let us assume first that the field particles form a heat bath. That is, they all stay together in the same state, despite the presence of the Brownian particle, because their response to the Brownian particle is very quick and because there are very many of them. In comparison, the Brownian particle moves slowly in the "fluid" of these field particles and encounters resistance. Our aim is to derive a molecular expression for such resistance (*3–5a*).

The slowness of the Brownian particle is reflected in the magnitude of the operator iL_B. For thermal motion, we expect $\mathbf{P} \sim (MkT)^{1/2}$. Thus, iL_B should be of order $M^{-1/2}$ in contrast to $iL \sim m^{-1/2}$, where the mass M of the Brownian particle is being considered much larger than m, the mass of the field particles. For this reason, the integral equation may be solved by perturbation.

The first approximation to $f^{(N+1)}$ may be given in a product form,

$$f^{(N+1)}(s) = f_0^{(N/1)} f_B(s), \tag{5.8}$$

where the conditional probability for the N field particles in the presence of the Brownian particle is replaced by its equilibrium expression:

$$f_0^{(N,1)} = \frac{1}{Z_N} \exp\left\{ -\frac{1}{kT} \left[\sum_i \frac{p_i^2}{2m} + \sum_{i>j} \Phi_{ij} + \sum_i \Phi(\mathbf{r}_i - \mathbf{R}) \right] \right\}. \tag{5.9}$$

Z_N is the partition function:

$$Z_N = \frac{1}{N!} \int \exp\left\{ -\frac{1}{kT} \left[\sum_i \frac{p_i^2}{2m} + \sum_{i>j} \Phi_{ij} + \sum_i \Phi(\mathbf{r}_i - \mathbf{R}) \right] \right\} d\{N\}. \tag{5.10}$$

The function appearing in the integrand of Eq. (5.6) is found as follows:

$$(-iL_B) f^{(N+1)} = \left(\frac{\mathbf{P} \cdot \mathbf{X}}{kT} f_B + \frac{\mathbf{P}}{M} \cdot \frac{\partial}{\partial \mathbf{R}} f_B + M\mathbf{X} \cdot \frac{\partial}{\partial \mathbf{P}} f_B \right) f_0^{(N/1)}. \tag{5.11}$$

The force $M\mathbf{X}$ acting on the Brownian particle is random and fluctuating. This random character is represented by

$$\int \mathbf{X} f_0^{(N/1)}\, d\{N\} = 0. \tag{5.12}$$

Owing to this equality, the second term of Eq. (5.11) gives no contribution in Eq. (5.7). On the other hand, the first and third terms, which are proportional to $\mathbf{X}$, may give contributions.

We seek the source of irreversibility in these contributions; we note that the operator $e^{-i(t-s)L}$ in Eq. (5.6) acts only on the fluctuating force $M\mathbf{X}$. This agrees with our previous assumption that the Brownian particle moves slowly. Thus, we introduce a new notation:

$$e^{i(t-s)L}\mathbf{X} = \mathbf{X}(t-s). \tag{5.13}$$

With these considerations, the right-hand side of Eq. (5.7) becomes

$$-\frac{\partial}{\partial \mathbf{P}} \cdot \int \frac{\partial \Phi(\mathbf{r}-\mathbf{R})}{\partial \mathbf{R}} \cdot \mathbf{X}(t-s)\left(\frac{\mathbf{P}}{kT} + M\frac{\partial}{\partial \mathbf{P}}\right) f^{(1+1)}(s)\, d\{1\}\, ds.$$

We define the time average of a physical quantity by

$$\bar{G}\bar{f}_{\mathrm{B}}(t) = \frac{1}{\tau}\int_0^\tau G(t+u) f_{\mathrm{B}}(t+u)\, du, \tag{5.14}$$

where

$$\bar{f}_{\mathrm{B}}(t) = \frac{1}{\tau}\int_0^\tau f_{\mathrm{B}}(t+u)\, du \tag{5.15}$$

is a time-smoothed distribution function. Then, performing such a time-smoothing operation on Eq. (5.7), we obtain

$$\frac{\partial \bar{f}_{\mathrm{B}}}{\partial t} + \frac{\mathbf{P}}{M}\cdot\frac{\partial}{\partial \mathbf{R}}\bar{f}_{\mathrm{B}} = -\frac{\partial}{\partial \mathbf{P}}\cdot\gamma\cdot\left(\mathbf{P} + MkT\frac{\partial}{\partial \mathbf{P}}\right)\bar{f}_{\mathrm{B}}, \tag{5.16}$$

where

$$\gamma = \frac{M}{kT\tau}\int_0^\tau\int_0^{t+u} \langle \mathbf{X}\mathbf{X}(t+u-s)\rangle\, ds\, du. \tag{5.17}$$

The tensor γ is determined by the time correlation of the random force on the Brownian particle. One may expect that the correlation decays very quickly with increasing time, and γ approaches a plateau value after a certain value of τ. Equation (5.16) is called the *Fokker–Planck equation*. The formal properties of this equation will be discussed in Section 8. Here, we only remark that the Fokker–Planck equation is reduced to the Langevin equation by integration over $\mathbf{P}$. Thus, we have arrived at our goal to obtain a molecular expression for the frictional constant (per unit mass).

In Eq. (5.17), $\langle\cdots\rangle$ represents the ensemble average in terms of $f_0^{(N/1)}$. Since this distribution function, $\mathbf{X}$, and the operator L do not include the momentum $\mathbf{P}$ of the Brownian particle, γ is independent of $\mathbf{P}$. Moreover, it should be a constant independent of t and τ so long as τ is sufficiently long. Thus, one can evaluate γ by replacing the double time integrals by a single time integral over infinite time. Using a new variable t instead of $t+u-s$, the friction constant per unit mass of a Brownian particle in isotropic fluids is given by

$$\gamma = \mathbf{1}\gamma,$$
$$\gamma = \frac{M}{3kT}\int_0^\infty \langle \mathbf{X}\cdot\mathbf{X}(t)\rangle\, dt, \tag{5.18}$$

where we have a scalar product of two $\mathbf{X}$ vectors and $\mathbf{1}$ is a unit tensor. The total friction constant

$$\zeta = M\gamma \tag{5.19}$$

is determined by the time average of the ensemble average of the correlation of the fluctuating force $M\mathbf{X}$ on the Brownian particle.

7.6. Autocorrelation Function

In the previous section, we saw that the friction coefficient is determined by the ensemble average of the correlation of the fluctuating force. In general, such correlation functions play important roles in nonequilibrium statistical mechanics. Therefore, let us study the properties of an autocorrelation function.

The *autocorrelation function* of a random variable $X(t)$ is defined by

$$G(s) = \langle X(t)X(t+s)\rangle, \tag{6.1}$$

where $\langle\cdots\rangle$ represents an equilibrium ensemble average. By a random variable, we mean $\langle X\rangle = 0$. Thus, the autocorrelation function is designed to give the dispersion $\langle X^2\rangle$ for $s = 0$. Of course, it is real for a real X. Even if X is complex, $G(s)$ can still be real if there is no correlation between the real and the complex parts.

Note that the autocorrelation function is independent of t, since it is defined in the equilibrium ensemble. Thus, by changing t to $t-s$ in Eq. (6.1), we find the symmetry property

$$G(s) = G(-s). \tag{6.2}$$

The autocorrelation function $G(s)$ has a maximum for $s = 0$. This property can be seen from

$$\langle [X(t) \pm X(t+s)]^2\rangle = \langle X^2(t)\rangle + \langle X^2(t+s)\rangle \pm 2\langle X(t)X(t+s)\rangle \geqslant 0,$$

or

$$G(0) \geqslant G(s). \tag{6.3}$$

We may assume that there is no correlation between $X(t)$ and $X(t+s)$ for a large time interval s. because X is a random variable. Thus, $G(s)$ decays to zero for large s:

$$\lim_{s\to\infty} G(s) = 0. \tag{6.4}$$

The ensemble average in the definition of an autocorrelation function may be replaced by a time average if the ergodic theorem is applicable to the product $X(t)X(t+s)$:

$$G(s) = \lim_{\tau\to\infty} \frac{1}{\tau}\int_{-\tau/2}^{\tau/2} X(t)X(t+s)\,dt. \tag{6.5}$$

In view of the random character of X and also of property (6.4), we introduce Fourier transforms:

$$\begin{aligned} X(t) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{X}(\omega)e^{i\omega t}\,d\omega; \\ G(s) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{G}(\omega)e^{-i\omega s}\,d\omega. \end{aligned} \tag{6.6}$$

From Eq. (6.5), we find

$$\tilde{G}(\omega) = \frac{1}{\tau}\,\tilde{X}(\omega)\,\tilde{X}(-\omega). \tag{6.7}$$

Thus,

$$\tilde{G}(\omega) = \tilde{G}(-\omega), \tag{6.8}$$

and the second of Eqs. (6.6) becomes

$$G(s) = \frac{1}{2\pi}\int_0^{\infty} 2\tilde{G}(\omega)\cos\omega s\,d\omega. \tag{6.9}$$

The function

$$D(\omega) = 2\tilde{G}(\omega) \tag{6.10}$$

is called the *spectral density* of the autocorrelation function. Thus, Eq. (6.9) states that the autocorrelation function is the Fourier cosine transform of its spectral density. This result is called the *Wiener–Khinchin theorem.*

Equations (6.9) and (6.10), together with (6.5) give

$$\begin{aligned} G(0) &= \lim \frac{1}{\tau}\int_{-\tau/2}^{\tau/2} X^2(t)\,dt \\ &= \frac{1}{2\pi}\int_0^{\infty} D(\omega)\,d\omega. \end{aligned} \tag{6.11}$$

Also, integrating Eq. (6.9) over s, we obtain

$$\begin{aligned} \int_0^{\infty} G(s)\,ds &= \int_0^{\infty} \langle X(t)X(t+s)\rangle\,ds \\ &= \tfrac{1}{4}\lim_{\omega\to 0+} D(\omega). \end{aligned} \tag{6.12}$$

Based on these general considerations of the autocorrelation function, let us now treat the particular case of the Langevin equation for one-dimensional Brownian motion. Using Eq. (4.3), we obtain

$$\langle v^2\rangle = \langle v^2(0)\rangle e^{-2\gamma t} + e^{-2\gamma t}\int_0^t\int_0^t e^{\gamma(x+y)}X(x)X(y)\,dx\,dy. \tag{6.13}$$

On the right-hand side of this equation, we find an autocorrelation function. As in the general case, we assume that the correlation function depends only on the difference of the time variables. Then, it is convenient to introduce new integration variables such that

$$u = x + y, \qquad s = x - y, \tag{6.14}$$

and perform the u-integration first for a given s. Since $x = s + y$ and the minimum of y is zero, the lower boundary of u is s. The maximum of y is $(\max x) - s = t - s$. Thus, the upper bound of u for a given s is $2t - s$. Accordingly, we can perform the double integrals as follows:

$$\begin{aligned} \int_0^t\int_0^t e^{\gamma(x+y)}G(x-y)\,dx\,dy &= \frac{1}{2}\,2\int_0^t G(s)\,ds\int_s^{2t-s} e^{\gamma u}\,du \\ &= \frac{1}{\gamma}\int_0^t ds\,G(s)(e^{\gamma(2t-s)} - e^{\gamma s}). \end{aligned} \tag{6.15}$$

Let us consider the limit $t\to\infty$ in Eq. (6.13). Due to the exponential function outside of the double integrals, only the first term on the right-hand side of Eq. (6.15) contributes. Making use of Eq. (6.9), we obtain

$$\begin{aligned} \langle v^2\rangle &= \frac{1}{\gamma}\int_0^{\infty} d\omega\,\frac{1}{2\pi}\int_0^{\infty} ds\,D(\omega)(\cos\omega s)e^{-\gamma s} \\ &= \frac{1}{2\pi}\int_0^{\infty} d\omega\,\frac{D(\omega)}{\omega^2+\gamma^2}. \end{aligned} \tag{6.16}$$

According to the equipartition law, the left-hand side is equal to kT/m. Thus,

$$\frac{kT}{m} = \frac{1}{2\pi}\int_0^{\infty} d\omega\,\frac{D(\omega)}{\omega^2+\gamma^2}. \tag{6.17}$$

In the above discussion, the spectral density has not been specified. When the time correlation between the X's exists only for a very short time interval,

$G(s)$ will behave like a delta function. In such a case, we may assume a white spectrum:

$$\begin{aligned} D(\omega) &= D(0+) \\ &= D, \end{aligned} \tag{6.18}$$

where D is a constant and $0+$ represents the limit $\omega \to 0$ from positive ω. Integrating Eq. (6.17) and making use of Eq. (6.12), we obtain

$$\gamma = \frac{m}{kT} \int_0^\infty \langle X(t)X(t+s)\rangle \, ds. \tag{6.19}$$

This is the one-dimensional version of Eq. (5.18).

7.7. Neutron Scattering

In scattering experiments on condensed matter, the effectiveness of probing radiation depends on its energy and momentum. When these quantities are of the right orders of magnitude, a maximum efficiency is expected. Electron beams and electromagnetic beams usually have energies of the order of 10 keV, while neutron beams cover the range of 1–100 MeV. The wavelength of the radiation must be comparable to the spacing of the atoms. For liquids and solids, 2π divided by the spacing is of the order of several inverse Ångströms.

In discussing the efficiency of a radiation probe, a quantum relation

$$\hbar\omega = \frac{\hbar^2 k^2}{2m}$$

between the energy and the momentum enters. Because of this form, it is generally not possible to choose both ω and k of the right orders of magnitude. However, for neutrons, this energy–momentum relation gives the wavelengths of atomic dimensions and the thermal energies. For this reason, neutron scattering experiments have become important for condensed matter. On the other hand, electron beams should have rather high energies in order to make the Born approximation applicable, and then the wavelength becomes relatively small. In such a case, only low-angle scattering becomes effective. In general, X rays and Mössbauer γ-rays are used only for spacing (i.e., momentum) analyses.

The scattering of electromagnetic waves is determined by the spatial correlation of the scattering particles. For neutron scattering, however, not only the spatial correlation, but also the time correlation of the scattering particles comes in.

The space–time correlation function introduced by Van Hove in 1954 (*6*) is defined to give the probability of finding a particle at time t at a position $\mathbf{r}$ from the origin, where there was a particle when $t = 0$. Since the density operator at t and $\mathbf{r}$ in an N-particle fluid is given in terms of the Dirac delta function by

$$n_t(\mathbf{r}) = \sum_{i=1}^{N} \delta(\mathbf{r} - \mathbf{r}_i(t)),$$

this probability may be expressed by

$$G(\mathbf{r}, t) = \frac{1}{N} < \sum_{i,j} \int d\mathbf{r}' \langle \delta[\mathbf{r} + \mathbf{r}_i(0) - \mathbf{r}'] \, \delta[\mathbf{r}' - \mathbf{r}_j(t)]\rangle, \tag{7.1}$$

where

$$\mathbf{r}_j(t) = e^{iHt/\hbar} \mathbf{r}_j \, e^{-iHt/\hbar} \tag{7.2}$$

is the position operator of the jth particle in the Heisenberg representation, H is the Hamiltonian of the system without the neutrons, and $\langle \cdots \rangle$ represents the statistical average over the ensemble for the system.

In classical systems, the vectors $\mathbf{r}_j(t)$ and $\mathbf{r}_i(0)$ are commuting quantities. In quantum systems, this is generally no longer the case for $t \neq 0$, and $G(\mathbf{r}, t)$ becomes complex. We shall write

$$G(\mathbf{r}, t) = G_0(\mathbf{r}, t) + iG_1(\mathbf{r}, t), \tag{7.3}$$

where G_0 and G_1 are real. It can be shown that G_0 is even and G_1 is odd in their arguments. For $t = 0$, the function G_1 in the imaginary part vanishes, so that

$$\begin{aligned} G(\mathbf{r}, 0) = G_0(\mathbf{r}, 0) &= \delta(\mathbf{r}) + N^{-1} \left\langle \sum_{i \neq j} \delta(\mathbf{r} + \mathbf{r}_i - \mathbf{r}_j) \right\rangle \\ &= \delta(\mathbf{r}) + n^{-1} \rho_2(r), \end{aligned} \tag{7.4}$$

according to Eq. (7.19) of Chapter 6, where $\rho_2(r)$ is the p.d.f. when the last term corresponds to classical systems. For homogeneous fluids, $G(\mathbf{r}, t)$ gives the number density for large $\mathbf{r}$ and t.

The function $G(\mathbf{r}, t)$ is related to the differential angular and energy cross section of scattered neutrons as follows:

$$\frac{d^2\sigma}{d\Omega \, d\omega} = b^2 \frac{k}{k_0} S(q, \omega), \tag{7.5}$$

where

$$S(q, \omega) = \int \{\exp[i(\mathbf{q} \cdot \mathbf{r} - \omega t)]\}[G(\mathbf{r}, t) - n] \, d\mathbf{r} \, dt. \tag{7.6}$$

Equation (7.5) corresponds to the scattering of incident neutrons with a change of the wave vector from $\mathbf{k}_0$ to $\mathbf{k} = \mathbf{k}_0 - \mathbf{q}$ with the energy transfer $\hbar\omega$, where

$$\omega = \frac{\hbar(k_0^2 - k^2)}{2m}. \tag{7.7}$$

The factor k/k_0 accounts for the conversion from the neutron density to flux. The constant b in Eq. (7.5) is proportional to the total target area:

$$b^2 = \frac{a^2 N}{2\pi}. \tag{7.8}$$

We are considering spin-independent scattering by scatterers of the same scattering length a. In slow neutron scattering, the energy change of the neutrons is of the same order as the incident neutron energy, but the Born approximation is still applicable as in the case of X-ray scattering. Usually, the target nucleus is represented by the pseudopotential:

$$\phi(r) = \left(\frac{4\pi a\hbar^2}{m}\right)\delta(\mathbf{r}). \tag{7.9}$$

In the case of scattering of electromagnetic waves, one arrives at essentially the same formula as Eq. (7.5) under the following assumptions:

(*i*) The electrons causing the scattering move rigidly with the nucleus.
(*ii*) The electronic states are not changed in the scattering process.
(*iii*) The energy transfer is small.

According to Landau and Lifshitz (*6a*), the coefficient b for light scattering can be related to the frequency-dependent dielectric constant ε in the limit of small values of wave number as follows:

$$b = \frac{2}{3}\frac{k_0^2}{4\pi}\left[\frac{\partial\varepsilon(\omega_0)}{\partial n}\right]_T,$$

where ω_0 is the frequency and k_0 is the wave number of the incident light. This formula corresponds to the case when the incident and scattered light are polarized with the electric vector perpendicular to the $(\mathbf{k}_0, \mathbf{k})$ plane.

The function G defined by Eq. (7.1) is related to the density correlation function in the form

$$n_0 G(\mathbf{r}' - \mathbf{r}, t' - t) = \langle n_t(\mathbf{r}) n_{t'}(\mathbf{r}')\rangle, \tag{7.10}$$

where n_0 is the equilibrium density. This relation, when split into real and imaginary parts, yields

$$n_0 G_0(\mathbf{r}' - \mathbf{r}, t' - t) = \tfrac{1}{2}\langle n_t(\mathbf{r}) n_{t'}(\mathbf{r}') + n_{t'}(r') n_t(\mathbf{r})\rangle; \tag{7.11}$$

$$\mathrm{i}n_0 G_1(\mathbf{r}' - \mathbf{r}, t' - t) = \tfrac{1}{2}\langle n_t(\mathbf{r}) n_{t'}(\mathbf{r}') - n_{t'}(\mathbf{r}') n_t(\mathbf{r})\rangle. \tag{7.12}$$

One can show that G_1 is related to the local density disturbance induced by the neutrons. By calculating the density of the total system, that is, a neutron plus targets, at the time t and at $\mathbf{r}$ to first order in the scattering length a, we find the relation

$$n(\mathbf{r}, t) = n_0 - \left(\frac{4\pi a\hbar^2}{m}\right)\left(\frac{2n_0}{\hbar}\right)\int_{-\infty}^{t} dt' G_1(\mathbf{r} - \mathbf{r}_{t'}, t - t'). \tag{7.13}$$

We shall not give a proof of this result. However, Eq. (7.13) can be understood from the general consideration in Chapter 13 of the linear response of a system to external perturbations.

In thermal equilibrium, the numbers of particles leaving and entering a certain energy region have to be the same. Applying the condition of detailed balance, one can derive another important relation between G_0 and G_1. For this purpose, let us note that the process $\mathbf{k}_0 \to \mathbf{k}$ and its inverse $\mathbf{k} \to \mathbf{k}_0$ may be balanced as follows (*7*):

$$\frac{\sigma(\mathbf{k}_0 \to \mathbf{k})}{k^2}\frac{k_0^2}{\sigma(\mathbf{k} \to \mathbf{k}_0)} = e^{-\hbar\omega/kT}. \tag{7.14}$$

In terms of $S(q, \omega)$, this equation gives

$$S(q, -\omega) = S(q, \omega)e^{-\hbar\omega/kT}. \tag{7.15}$$

Thus,

$$\frac{S(q, \omega) - S(q, -\omega)}{2} = \tanh\left(\frac{\hbar\omega}{2kT}\right)\frac{S(q, \omega) + S(q, -\omega)}{2}. \tag{7.16}$$

Equation (7.6) yields

$$G_1(\mathbf{r}, t) = -\tanh\left(\frac{\hbar}{2kT}\frac{\partial}{\partial t}\right) G_0(\mathbf{r}, t). \tag{7.17}$$

In the classical limit, this relation results in

$$\frac{2kT}{\hbar} G_1(\mathbf{r}, t) = -\frac{\partial}{\partial t} G_0(\mathbf{r}, t). \tag{7.18}$$

7.8. The Fokker–Planck Equation

The probability distribution function $f_2(x_2, t_2; x_1, t_1)$ of two sets of variables (x_1, t_1) and (x_2, t_2) may be related to a singlet distribution function $f_1(x_1, t_1)$ in the form

$$f_2(x_2, t_2; x_1, t_1) = K(x_2, t_2; x_1, t_1) f_1(x_1, t_1), \tag{8.1}$$

where K is the conditional probability for the set (x_2, t_2) for a given (x_1, t_1). Since f_2 and f_1 are positive, K is also positive. Also, if

$$f_1(x_2, t_2) = \int f_2(x_2, t_2 ; x_1, t_1)\, dx_1, \tag{8.2}$$

the following relations hold:

$$\int K(x_2, t_2 ; x_1, t_1)\, dx_2 = 1, \tag{8.3}$$

$$\int K(x_2, t_2 ; x_1, t_1) f_1(x_1, t_1)\, dx_1 = f_1(x_2, t_2). \tag{8.4}$$

If the conditional probability or the transition probability K from (x_1, t_1) to (x_2, t_2) depends only on these two points, the process is called a *Markov process*. If the transition depends on the processes prior to (x_1, t_1), the process is non-Markovian. The series of events $x_1, x_2, \ldots$ evolved by K forms a *Markov chain*. As one can see easily by considering a series of processes $(x_1, t_1) \to (x_2, t_2) \to (x_3, t_3)$, the transition probability satisfies the *Chapman–Kolmogorov equation*:

$$K(x_3, t_3 ; x_1, t_1) = \int K(x_3, t_3; x_2, t_2) K(x_2, t_2 ; x_1, t_1)\, dx_2, \tag{8.5}$$

that is, the transition probability $K(x_3, t_3 ; x_1, t_1)$ is obtained by considering all the transitions involving the intermediate points x_2.

In Eqs. (8.4) and (8.5), no integration over t is necessary because the distribution functions f_1 and f_2 are both normalized in the x-space. Note that the symbol x may represent many variables determined by the same t. Thus, $x = x(t)$, and the transition probability K may be considered to be dependent on the time difference $t_2 - t_1$.

If a particle moves in phase space from $x = (\mathbf{r}, \mathbf{p} - \Delta\mathbf{p})$ at t to $x + \Delta x = (\mathbf{r} + \Delta\mathbf{r}, \mathbf{p})$ at $t + \Delta t$, Eq. (8.4) reads for $f = f_1$

$$f(\mathbf{r} + \Delta\mathbf{r}, \mathbf{p} ; t + \Delta t) = \int K(t + \Delta t ; t) f(\mathbf{r}, \mathbf{p} - \Delta\mathbf{p} ; t)\, d(\Delta\mathbf{r})(d\Delta\mathbf{p}), \tag{8.6}$$

where we have abbreviated the arguments in K. For small Δt, Newton's mechanics is applicable. Thus, $\Delta\mathbf{r} = \mathbf{p}\,\Delta t/m$, and we put

$$K(t + \Delta t ; t) = W(\Delta\mathbf{p} ; \mathbf{p})\, \delta\left(\Delta\mathbf{r} - \frac{\mathbf{p}\Delta t}{m}\right). \tag{8.7}$$

where $W(\Delta\mathbf{p} ; \mathbf{p})$ represents the transition probability from $\mathbf{p}$ to $\mathbf{p} + \Delta\mathbf{p}$.

Introducing Eq. (8.7) into Eq. (8.6) and expanding both sides in $\Delta\mathbf{r}$, $\Delta\mathbf{p}$, and Δt, we obtain

$$\left(\frac{\partial f}{\partial t} + \frac{\mathbf{p}}{m} \cdot \nabla_{\mathbf{r}} f\right) \Delta t + 0(\Delta t^2) = -\nabla_{\mathbf{p}} \cdot \langle \Delta\mathbf{p} \rangle f + \frac{1}{2} \nabla_{\mathbf{p}} \nabla_{\mathbf{p}} : (\langle \Delta\mathbf{p}\, \Delta\mathbf{p} \rangle f) + \cdots. \tag{8.8}$$

Thus, we arrive at the Fokker–Planck equation in phase space:

$$\frac{\partial f}{\partial t} + \frac{\mathbf{p}}{m} \cdot \nabla_{\mathbf{r}} f = -\nabla_{\mathbf{p}} \cdot \frac{\langle \Delta\mathbf{p} \rangle}{\Delta t} f + \frac{1}{2} \nabla_{\mathbf{p}} \nabla_{\mathbf{p}} : \frac{\langle \Delta\mathbf{p}\, \Delta\mathbf{p} \rangle}{\Delta t} f + \cdots. \tag{8.9}$$

This type of differential equation can be solved for a suitable boundary condition with the aid of the characteristic differential equations that represent the motion of the particle. As in the case of the Langevin equation, the motion of a particle in phase space may be considered to form a random walk process. In such a case, we assume that

$$\Delta\mathbf{p} = -\gamma\mathbf{p}\, \Delta t + m\mathbf{X}\, \Delta t, \tag{8.10}$$

where $m\mathbf{X}$ is a fluctuating force, and that the transition probability $W(\Delta\mathbf{p} ; \mathbf{p})$ is of the form of $\tau\,(\overline{\mathbf{X}}(\Delta t))$ of Eq. (4.5):

$$W(\Delta\mathbf{p} ; \mathbf{p}) = \frac{1}{(4\pi m^2 q\, \Delta t)^{3/2}} \exp\left[\frac{-|\Delta\mathbf{p} + \gamma\mathbf{p}\, \Delta t|^2}{4m^2 q\, \Delta t}\right]. \tag{8.11}$$

Then, on the average, we obtain

$$\langle \Delta\mathbf{p} \rangle = -\gamma\mathbf{p}\, \Delta t, \qquad \langle \Delta p_i\, \Delta p_j \rangle = 2m^2 q\, \Delta t\, \delta_{ij}. \tag{8.12}$$

Therefore, for an isotropic case, Eq. (8.9) is brought into

$$\frac{\partial f}{\partial t} + \frac{\mathbf{p}}{m} \cdot \nabla_{\mathbf{r}} f = \gamma \nabla_{\mathbf{p}} \cdot (f\mathbf{p}) + q m^2 \nabla_{\mathbf{p}}{}^2 f, \tag{8.13}$$

that is, the Fokker–Planck equation in standard form.

The solution for this type of equation, which reduces to the Maxwell distribution for equilibrium, may be obtained easily by using the characteristic equations, which describe the motion of the particle. The solution is similar to that of the Langevin equation in coordinate space.

The Fokker–Planck equation manifests irreversibility. The two terms on the right-hand side represent, respectively, a frictional and a diffusional change. Thus, this indicates that, by considering diffusion in momentum space, one can obtain the irreversible coefficient (δ).

7.9.* Self-Avoiding Walk Problem

The theory of random walks has been applied successfully to polymer and other problems. However, interactions existing in real polymer molecules cause phase transitions and other important effects. Even if we adopt a simple pearl-necklace model in which a polymer chain consists of hard-sphere

units with a constant bond length, we face the *self-avoiding walk problem*. In this problem, an otherwise random walker walks about under the condition that he does not step on his footprints. Then, we can expect that the walker reaches farther than in the case of purely random walks. Correspondingly, the polymer chain will be expanded. Let us find such an expansion effect in the square average end-to-end distance of a flexible chain molecule constituting N units.

For our purpose, we introduce the partition function $Z(r)$ of a chain of a given end-to-end distance r, given by

$$Z(r) = \int \prod_{i=1}^{N} d\mathbf{r}_i \times \prod_{i=1}^{N-1} \delta(|\mathbf{r}_{i+1} - \mathbf{r}_i| - b)\, \delta(|\mathbf{r}_1|)(4\pi b^2)^{-N}\, \delta(\mathbf{r} - \mathbf{r}_N) e^{-\beta\Phi}, \tag{9.1}$$

where $\Phi = \sum \phi_{ij}$ is the total intramolecular potential, b is the bond length, and $\mathbf{r}$ is a given vector which connects the two ends of the chain, one end being fixed at the origin of the coordinates. The Dirac δ-functions restrict the distance between the neighboring units, and $\mathbf{r}_i$ represents the coordinates of the ith unit.

The Boltzmann factor in the integrand may be expanded in a series in terms of the Mayer f-function $f(r) = e^{-\beta\phi(r)} - 1$. Introducing an order parameter ε to indicate the number of f-functions, we can express Eq. (9.1) in the form

$$Z(r, \varepsilon) = \sum_l \frac{\varepsilon^l G_l(r)}{l!}. \tag{9.2}$$

It is convenient to introduce a transformation

$$\tilde{Z}(\lambda, \varepsilon) = \int (\exp - \lambda r^2) Z(r, \varepsilon)\, d\mathbf{r}. \tag{9.3}$$

Then, from Eq. (9.2), we obtain (9)

$$\tilde{Z}(\lambda, \varepsilon) = \sum_l \frac{\varepsilon^l g_l}{l!}, \qquad g_l = \int G_l(r)(\exp - \lambda r^2)\, d\mathbf{r}. \tag{9.4}$$

As we have discussed in Section 5.5, the logarithm of $\tilde{Z}$ may be expanded in powers of ε with the Thiele semiinvariants as the coefficients:

$$\log \tilde{Z} = \sum_l \frac{\varepsilon^l b_l(\lambda)}{l!}, \tag{9.5}$$

$$\begin{aligned} b_0 &= \log g_0, \\ b_1 &= \frac{g_1}{g_0}, \\ b_2 &= \frac{g_2}{g_0} - \left(\frac{g_1}{g_0}\right)^2, \\ b_3 &= \frac{g_3}{g_0} - \frac{3g_1 g_2}{g_0^2} + 2\left(\frac{g_1}{g_0}\right)^3, \\ &\cdots. \end{aligned} \tag{9.6}$$

The formal integral transform of Eq. (9.3) gives a cluster expansion for the average square end-to-end distance r^2 as follows:

$$\langle r^2 \rangle = \sum_l \frac{1}{l!} \left[\frac{\partial b_l(\lambda)}{\partial(-\lambda)}\right]_{\lambda=0}. \tag{9.7}$$

Thus, $\langle r^2 \rangle$ can be obtained by the same operation as for the pressure of an imperfect gas.

To go further, it is convenient to introduce bond vectors by

$$\begin{aligned} &\mathbf{s}_i = \mathbf{r}_i - \mathbf{r}_{i-1}, \qquad i = 2, 3, \ldots, N, \\ &\mathbf{s}_1 = \mathbf{r}_1 = 0. \end{aligned} \tag{9.8}$$

The partition function $Z(r)$ becomes

$$Z(r) = \int \prod_{i=2}^{N} d\mathbf{s}_i\, \delta(|\mathbf{s}_i| - b)(4\pi b^2)^{-N}\, \delta\left(\mathbf{r} - \sum_{i=2}^{N} \mathbf{s}_i\right) \times \prod_{l<m} [1 + f(\mathbf{s}_{l+1} + \cdots + \mathbf{s}_m)]. \tag{9.9}$$

The first coefficient in the expansion of Eq. (9.2) can be reduced to

$$G_1(r) = \sum_{l<m} \int \prod_{i=2} d\mathbf{s}_i\, \delta(|\mathbf{s}_i| - b)(4\pi b^2)^{-N} \times \delta\left(\mathbf{r} - \left(\sum_2^l \mathbf{s}_i + \sum_{l+1}^m \mathbf{s}_i + \sum_{m+1}^N \mathbf{s}_i\right)\right) f\left(\sum_{l+1}^m \mathbf{s}_i\right). \tag{9.10}$$

Thus, it is convenient to introduce new variables by

$$\mathbf{x} = \sum_2^l \mathbf{s}_i, \qquad \mathbf{y} = \sum_{l+1}^m \mathbf{s}_i, \qquad \mathbf{z} = \sum_{m+1}^N \mathbf{s}_i. \tag{9.11}$$

Then, we arrive at

$$G_1(r) = \sum_{l<m} \int G^0_{l-1}(x) G^0_{m-l}(y) G^0_{N-m}(z) f(\mathbf{y})\, \delta(\mathbf{r} - \mathbf{x} - \mathbf{y} - \mathbf{z})\, d\mathbf{x}\, d\mathbf{y}\, d\mathbf{z}, \tag{9.12}$$

where

$$G_l{}^0(x) = \frac{1}{(2\pi)^3} \int [S(K)^l] \exp i\mathbf{K} \cdot \mathbf{x}\, d\mathbf{K}, \tag{9.13}$$

$$S(K) = \frac{\sin Kb}{Kb}. \tag{9.14}$$

The meaning of Eq. (9.12) is clear. It simply states that $G_1(r)$ is the sum of single-loop diagrams such as illustrated in Fig. 7.1. Here, the solid line

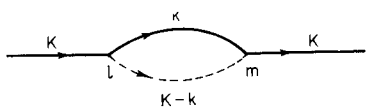


Fig. 7.1. Momentum-space representation of first-order correlation. The momentum representing the propagation of the chain is conserved at vertices l and m, where an interaction line leaves and enters.

represents the chain and the dotted line $f(\mathbf{y})$. Note that the graph shows the integrand of Eq. (9.12) in momentum space. Using Eq. (9.13) and performing the coordinate-space integration, we find

$$G_1(r) = \sum_{l<m} \frac{1}{(2\pi)^6} \int (\exp i\mathbf{K} \cdot \mathbf{r}) [S(K)]^{N-1} \left[\frac{S(k)}{S(K)}\right]^{m-l} \times \tilde{f}(\mathbf{K} - \mathbf{k})\, d\mathbf{K}\, d\mathbf{k}, \tag{9.15}$$

where $\tilde{f}(k)$ is the Fourier transform of the Mayer f-function.

Note that in Eq. (9.15) only two momenta $\mathbf{K}$ and $\mathbf{k}$ appear, as a result of the momentum conservation, which is illustrated in Fig. 7.1. The power $(N-1)$ of the first $S(K)$ in Eq. (9.15) corresponds to the total number of steps and the power $(m-l)$ of $[S(k)/S(K)]$ to the number in the loop. All higher-order coefficients can be represented by similar self-energy-type graphs, and analogous simple rules hold for constructing their integrands. Once a configuration of the chain is specified, its contribution to $Z(r)$ can easily be written down.

For large r, the main contribution to the integral of Eq. (9.13) comes from a small-$\mathbf{K}$ region, particularly for large l. Thus, by expanding $S(K)$, we find an asymptotic expression

$$G_l{}^0(x) = \left(\frac{3}{2\pi b^2 l}\right)^{3/2} \exp\left(\frac{-3r^2}{2b^2 l}\right). \tag{9.16}$$

From Eq. (9.7), we obtain the first approximation for $\langle r^2 \rangle$ for a long chain in the form

$$\langle r^2 \rangle = \langle r^2 \rangle^0 + \tfrac{1}{2} N^2 \{\langle r^2 f(r) \rangle^0 - \langle r^2 \rangle^0 \langle f(r) \rangle^0\}, \tag{9.17}$$

where $\langle \cdots \rangle^0$ represents the average taken with the probability function $G_N{}^0(r)$.

For more explicit results, one needs to specify the potential function $\phi(r)$. For a pearl-necklace model, it is easy to find

$$\tilde{f}(k) = -\frac{4\pi}{k^3} (\sin ak - ak \cos ak), \tag{9.18}$$

where a is the hard-sphere diameter. Since $\tilde{f}(k)$ is small except for small values of k, we may approximate it by $-4\pi a^3/3$. For a long chain, we then obtain

$$\langle r^2 \rangle = Nb^2 + \frac{8}{(6\pi)^{1/2}} \frac{a^3}{b} N^{3/2}, \tag{9.19}$$

where the first term is the random walk result and the second term represents the first-order self-avoiding walk effect. As we expected, this term is positive, so that the chain is expanded on the average.

One can evaluate higher-order terms in a similar way. However, the calculation becomes increasingly difficult. Indeed, it is still an open question to find $\langle r^2 \rangle$ for all orders.

Problems

1. The diffusion equation

$$\frac{\partial W}{\partial t} = D \nabla^2 W$$

can be solved in terms of the Green's function $K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1)$ satisfying

$$\frac{\partial K}{\partial t_2} - D \nabla_2{}^2 K(\mathbf{r}_2, t_2; r_1, t_1) = \delta(t_2 - t_1)\, \delta(\mathbf{r}_2 - \mathbf{r}_1),$$

where ∇_2 operates on the coordinates with index 2. Assuming that the Brownian particle was at $\mathbf{r} = 0$ at $t = 0$, show that

$$K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) = \frac{\theta(t_2 - t_1)}{[4\pi D(t_2 - t_1)]^{3/2}} \exp\left[-\frac{(\mathbf{r}_2 - \mathbf{r}_1)^2}{4D(t_2 - t_1)}\right],$$

where

$$\theta(t) = \begin{cases} 1, & t > 0, \\ 0, & t < 0. \end{cases}$$

2. Show that, in the above problem,

$$K(\mathbf{r}_2, t_2; \mathbf{r}_1, t_1) = \int K(\mathbf{r}_2, t_2; r_3, t_3) K(\mathbf{r}_3, t_3; \mathbf{r}_1, t_1)\, \mathbf{dr}_3.$$

3. Determine in integral form the probability for a Brownian particle to pass through n gates which restrict the trajectories such that

$$\mathbf{a}_j \leqslant \mathbf{r}_j \leqslant \mathbf{b}_j, \qquad j = 1, 2, \ldots, n,$$

where the $\mathbf{a}_j$ and $\mathbf{b}_j$ are given constant vectors.

4. A functional $F[\mathbf{r}(t)]$ of the trajectory of a Brownian particle may be considered as the limit of a function of $n + 1$ points $\mathbf{r}_0, \mathbf{r}_1, \ldots, \mathbf{r}_n$, where $\mathbf{r}_j = \mathbf{r}[j(t - t_0)/n]$ and $\mathbf{r}_0 = \mathbf{r}(0)$. Write down the expression for the average of the functional over the paths of the Brownian particle that started moving from $\mathbf{r}_0$ at time 0.

5. The diffusion of a Brownian particle in a field of force $\Phi(r)$ may be described by the differential equation

$$\frac{\partial W}{\partial t} = D\left[\nabla^2 W - \frac{\gamma}{kTD}\Phi(r) W\right].$$

Find the Green's function for this type of differential equation.

6. Using the random walk results of Eqs. (4.11) and (4.15), $\langle r^2 \rangle = 6Dt$, and observing that the displacement vector $\mathbf{r}$ is related to the velocity by

$$\mathbf{r}(t) = \int_0^t \mathbf{v}(t')\, dt',$$

express the diffusion constant D in terms of the velocity correlation function. Assume that the velocity correlation depends only on the time difference and decays to zero fast enough so that the integration over the time difference may be extended to infinity.

7. For a charged particle, Eq. (4.12) may be rewritten as

$$D = \frac{kT}{e}\frac{e}{m\gamma} = \frac{kT}{e}\mu,$$

where μ is the mobility per unit electric field. Using this so-called *Nernst relation*, express the mobility in terms of the velocity autocorrelation function. Also, consider expressing $e\mu$ in terms of $j = ev$ and interpret the result.

GENERAL REFERENCES

S. Chandrasekhar, *Rev. Mod. Phys.* **15**, 1 (1943).
M. Lax, *Rev. Mod. Phys.* **32**, 25 (1960).

REFERENCES

1. A. Einstein, *Ann. Physik.* **17**, 549 (1905).
*1a.*E. W. Montroll, *Commun. Pure Appl. Math.* **5**, 415 (1952).
2. A. Isihara, *Adv. Polymer Sci.* **5**, 531 (1968).
3. J. L. Lebowitz and R. J. Rubin, *Phys. Rev.* **131**, 2381 (1963).
4. P. Resibois and H. T. Davis, *Physica* **30**, 1077 (1964).
5. P. Mazur and I. Oppenheim, *J. Phys. Soc. Japan Suppl.* **26**, 35 (1969).
*5a.*M. S. Green, *J. Chem. Phys.* **19**, 1036 (1951); **20**, 1281 (1952); **22**, 398 (1954).
6. L. Van Hove, *Phys. Rev.* **95**, 249 (1954); *Physica* **24**, 404 (1958).
*6a.*L. D. Landau and I. M. Lifshitz, "Electrodynamics of Continuous Media," Chapter 14. Pergamon Press, Oxford, 1960.
7. P. Schöfield, *Phys. Rev. Lett.* **4**, 249 (1960).
8. M. Toda, *J. Phys. Soc. Japan* **9**, 440 (1954).
9. A. Isihara, *Adv. Polymer Sci.* **7**, 449 (1971).

8

Lattice Statistics

The statistical theory of ferromagnetism, which started with the thesis of Ising (1), has become the core of theories of phase transitions. His model consists of a spin lattice, each spin taking on one of two values, $+\frac{1}{2}$ and $-\frac{1}{2}$, and interacting only with its nearest neighbors. This model, for its simplicity, has the advantage of wide applicability and mathematical tractability.

Ising treated only the one-dimensional spin lattice and did not actually obtain the spontaneous magnetization characteristic of ferromagnetism. The first consideration of the ferromagnetic properties in a two-dimensional Ising lattice was given by Peierls in 1936 (2). In 1941, Kramers and Wannier extended the Ising lattice theory by using the concept of a dual lattice (3). They determined the transition point rigorously, assuming that there existed a phase transition point. In 1944, Onsager opened a new door for the study of the statistical mechanics of phase transitions by solving the two-dimensional Ising lattice problem rigorously and proving the existence of a phase transition. This was a monumental event because, before Onsager, it was not known whether a rigorous theory could show a phase transition.

The Ising lattice statistics is very useful because of its many applications, both magnetic and nonmagnetic. These include ferromagnetism, antiferromagnetism, ferrimagnetism, binary mixtures, the lattice model of fluids, adsorption on surfaces, DNA "melting," etc. Therefore, the Ising lattice statistics is of central importance in statistical mechanics.

Nevertheless, it must be mentioned that even before the phase transition in the Ising lattice was known, Heisenberg developed an alternative model for ferromagnetism in 1928 (4). In the ***Heisenberg model****, each lattice site has the total magnetic moment of an atom, which may be represented by a three-dimensional spin operator* **S**. *The interactions between nearest-neighbor atoms are expressed by the Hamiltonian*

$$H = -\sum_{(ij)} J\mathbf{S}_i \cdot \mathbf{S}_j .$$

Therefore, for ferromagnetism, this model is more realistic than the Ising model. However, the exchange integral J for interatomic distances is usually negative (it is positive for electrons at intraatomic distances). Thus, the explanation of ferromagnetism with the Heisenberg model has to invoke the positiveness of the exchange energy. The Ising model is also based on the same assumption.

The complexity of the Heisenberg Hamiltonian makes its statistical treatment difficult. Although high- and low-temperature cases have been treated analytically, a rigorous treatment of a Heisenberg model is difficult even for a one-dimensional case.

In parallel with the progress in the spin-lattice theories, the statistical mechanics of the frequency spectrum of lattice vibrations and the theory of lattice defects have been developed by Montroll (5) and others (6–8). Primarily, these theories take into consideration only the interactions with the nearest-neighbors. The advantage of this simplification, i.e., ***nearest-neighbor statistics****, is that one can develop rigorous theories. In contrast to these theories, the classical Weiss theory of ferromagnetism is based on a self-consistent and long-range field.*

In this chapter, we shall adopt the nearest-neighbor interaction model. The total number of spins or atoms will be denoted by N.

8.1. One-Dimensional Lattice

Let us denote the spin on the ith site in a one-dimensional crystal by σ_i. The partition function is given by

$$Z = \sum_{\sigma_1 = \pm 1} \cdots \sum_{\sigma_N = \pm 1} \prod_i e^{K\sigma_i\sigma_{i+1}}, \tag{1.1}$$

where the product is over all the spin pairs and

$$K = \frac{J}{2kT},$$

with J a constant determined by the exchange energy for a given lattice distance, and the summations are over all possible σ values, that is, ± 1. The factor $\frac{1}{2}$ in K has been introduced so that σ takes on values ± 1.

Let us introduce a matrix P with the elements $P(\sigma, \sigma') = e^{K\sigma\sigma'}$. Since σ is either 1 or -1, we find

$$P(\sigma, \sigma') = e^{K\sigma\sigma'} = \cosh K + \sigma\sigma' \sinh K. \tag{1.2}$$

Here, the last equality is an identity. Using this identity, the product in Eq. (1.1) may be expanded in the σ's. The properties

$$\sigma_i^2 = 1, \qquad \sum_{\sigma_i = \pm 1} \sigma_i = 0 \tag{1.3}$$

simplify the expansion. First, we note that the terms with odd powers of σ's vanish. Only the terms with even powers of σ's remain. For instance, the factor $\sigma_1\sigma_2\sigma_2\sigma_3 \cdots \sigma_s\sigma_1$ contributes. Because of the first relation in Eqs. (1.3) and the summation over spin values, such a factor always contributes 2^N. As a result, we obtain

$$Z = (2 \cosh K)^N + (2 \sinh K)^N. \tag{1.4}$$

For large N, note that only the first term is important.

We can arrive at the same result by a matrix method. Let us introduce an eigenvalue problem:

$$\sum_{\sigma'} P(\sigma, \sigma')a(\sigma') = \lambda a(\sigma). \tag{1.5}$$

More explicitly,

$$\begin{bmatrix} e^K & e^{-K} \\ e^{-K} & e^K \end{bmatrix} \begin{bmatrix} a(1) \\ a(-1) \end{bmatrix} = \lambda \begin{bmatrix} a(1) \\ a(-1) \end{bmatrix}. \tag{1.6}$$

From the secular equation, we find

$$\begin{aligned} \lambda &= e^K \pm e^{-K} \\ &= 2 \cosh K \quad \text{or} \quad 2 \sinh K. \end{aligned} \tag{1.7}$$

Thus, we arrive at Eq. (1.4). In particular, for large N, we find

$$Z = \lambda_{\max}^N. \tag{1.8}$$

The above method is applicable even if a magnetic field is present. We need only to introduce

$$Q = \begin{bmatrix} e^C & 0 \\ 0 & e^{-C} \end{bmatrix}, \qquad C = \frac{H}{kT}, \tag{1.9}$$

and consider the eigenvalue problem of a matrix QP. We then find the maximum eigenvalue $\lambda_{\max}$ given by

$$\lambda_{\max} = e^K \cosh C + (e^{2K} \sinh^2 C + e^{-2K})^{1/2}. \tag{1.10}$$

Thus, the intensity of magnetization is given in the form

$$I = \frac{\sinh C}{(\sinh^2 C + e^{-4k})^{1/2}}. \tag{1.11}$$

The above results show that the partition function of a one-dimensional Ising lattice does not have a singularity and does not yield a phase transition.

Nevertheless, by introducing a long-range exponential interaction, Kac, Hemmer, and Uhlenbeck (*9*) obtained a phase transition in a one-dimensional gas. Also, for a polymer chain it has been shown that the one-dimensional Ising lattice model is useful to interpret experimental "phase transition" curves, as will be discussed in the next section.

Example. *Tonks Equation*

The configurational partition function Q of a one-dimensional hard-sphere gas may be expressed in the form

$$Q = \int_{(2N-1)a/2}^{L-(a/2)} dx_N \cdots \int_{3a/2}^{x_3 - a} dx_2 \int_{a/2}^{x_2 - a} dx_1, \tag{1.12}$$

where a is the hard-sphere diameter. By changing variables to $y_j = x_j - \frac{1}{2}(2j-1)a$, show that

$$Q = \frac{1}{N!}(L - Na)^N \qquad (N \gg 1), \tag{1.13}$$

and derive the pressure

$$p = \frac{kT}{l - a}, \tag{1.14}$$

where $l = L/N$ is the "volume" per particle.

8.2.* Helix-Coil Transition in Polypeptide and "Melting" of DNA

Some synthetic polypeptides and proteins take on the α-helix structure in the ordered state and a random coil structure in the disordered state. The most popular synthetic polypeptide, poly-γ-benzyl-L-glutamate, is helical in ethylene dichloride and is random in dichloroacetic acid. For proteins and nucleic acids, random coils are stable at high temperatures. For poly-L-glutamic acid in aqueous dioxane, the helix structure is stable at low pH and low temperatures.

The DNA molecules assume the double helix structure when four different types of bases are bonded. They can "melt" if the bonds between the four bases are broken. If the molecules are replaced by a one-dimensional Ising model (*10*), the bonded and nonbonded states of DNA may be considered to correspond to the spin-up and spin-down states, although the four different bases cannot be distinguished with this simple model.

In the case of polypeptide chains, the interactions of H in the N—H of an amide group with the oxygen in the fourth subsequent amide residue are crucial in the formation of the ordered state. However, let us consider the helix-coil transition by a simpler one-dimensional chain with nearest-neighbor interactions. Since it is conceivable that one bond-formation induces another in a cooperative way, we construct a model partition function according to the rules:

(1) Every unbonded segment (−) contributes a factor 1.
(2) Every bonded segment (+) after a bonded segment contributes a factor s.
(3) Every bonded segment (+) after an unbonded segment (−) yields a factor s^*.

Where, s and s^* are parameters representing the Boltzmann factors. The cooperative character is reflected in the condition $s^* \neq s$.

Since we have four combinations of the neighboring segments, we introduce

$$P = \begin{bmatrix} 1 & s^* \\ 1 & s \end{bmatrix}. \tag{2.1}$$

If the first segment is unbonded, the partition function for N amide segments is given by

$$Z(s^*, s) = (1, 0)P^{N-1}\begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{2.2}$$

If λ_0 and λ_1 are the eigenvalues of P and if λ_0 is the largest eigenvalue, we find

$$Z(s^*, s) \sim \lambda_0^{N-1}. \tag{2.3}$$

λ_0 changes strongly near $s \sim 1$. Indeed, when $(1-s)^2 \gg 4s^*$, $\lambda_0 = 1 + (s^*/s)^{1/2}$ for $s = 1$ and

$$\begin{aligned} \lambda_0 &\sim 1 + \frac{2s^*}{1-s}, \qquad s < 1, \\ &\sim s + \frac{2s^*}{1-s}, \qquad s > 1. \end{aligned} \tag{2.4}$$

Thus, the average number of helical segments, given by

$$\langle n \rangle = \frac{\partial(\log Z)}{\partial(\log s)} = \frac{Ns}{\lambda_0}\frac{\partial \lambda_0}{\partial s}, \tag{2.5}$$

changes abruptly around $s = 1$, and we may consider $s = 1$ as corresponding to the helix-coil or DNA transitions.

8.3. Duality Principle

The partition function for a two-dimensional isotropic lattice is

$$Z = \sum_{\sigma_1 = \pm 1} \cdots \sum_{\sigma_N = \pm 1} \prod_{(ij)} e^{K\sigma_i\sigma_j}, \tag{3.1}$$

where (ij) indicates nearest-neighbor spin pairs. To evaluate the right-hand side, it is helpful to use the identity (1.2), which can be rewritten in the form:

$$e^{K\sigma\sigma'} = (\cosh K)(1 + \sigma\sigma' \tanh K). \tag{3.2}$$

Let S and z be, respectively, the total number of spin pairs and the coordination number. Then, $S = zN/2$. Let us introduce a notation

$$x = \tanh K$$

and expand Z in a power series in x as follows (*11*):

$$\begin{aligned} Z &= (\cosh K)^S \sum_{\sigma_1 = \pm 1} \cdots \sum_{\sigma_N = \pm 1} \prod_{(ij)} (1 + x\sigma_i\sigma_j) \\ &= (\cosh K)^S \sum_{\sigma_1 = \pm 1} \cdots \sum_{\sigma_N = \pm 1} \\ &\quad \times \left[1 + x\sum_{(ij)}(\sigma_i\sigma_j) + x^2 \sum_{(ij)}\sum_{(kl)}(\sigma_i\sigma_j)(\sigma_k\sigma_l) + \cdots\right]. \end{aligned} \tag{3.3}$$

The right-hand side may be illustrated by diagrams (Fig. 8.1) by connecting interacting spin pairs. The summations over spin values are subject to Eq. (1.2):

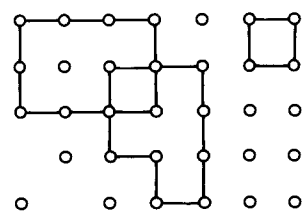


Fig. 8.1. Closed polygons in an Ising lattice.

$$\sum_i \sigma_i = 0, \qquad \sum_i \sigma_i^2 = 2. \tag{3.4}$$

Therefore, only the terms with even spin powers give nonvanishing contributions to Z. In terms of graphs, only closed polygons contribute to the partition function. Closed polygons with m interaction lines contribute $2^N x^m$ to the partition function irrespective of their shape. Therefore, the Ising statistics can be related to the probability that a random walker on a lattice returns to his starting point. If Ω_n is the number of closed polygons of n sides, the partition function assumes the form (*11*)

$$Z = 2^N(\cosh K)^S\left(1 + \sum_n \Omega_n x^n\right). \tag{3.5}$$

The same partition function can be obtained in a different way. Instead of counting Z in terms of closed polygons, let us use the number r of opposite spin pairs. Let S be the total number of spin pairs. Then,

$$\sum_{(ij)} \sigma_i \sigma_j = (S - r)(1) + r(-1) = S - 2r. \tag{3.6}$$

The partition function can be expressed in the form

$$Z = 2e^{SK}\left(1 + \sum_r \omega_r e^{-2Kr}\right), \tag{3.7}$$

where ω_r is the number of configurations of the lattice with r opposite spin pairs, and where the summation is over all possible r. The factor 2 is necessary because the direction of all the spins can be reversed for the same opposite-spin pairs.

We now introduce the concept of a *dual lattice* (*3*). For every two-dimensional lattice, we construct its dual lattice by drawing lines to bisect the lines connecting the nearest-neighbor lattice sites (Fig. 8.2). When the new lattice

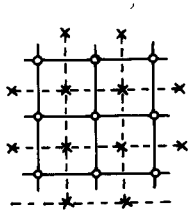


Fig. 8.2. A square lattice (solid lines) and the dual lattice (dashed lines) are the same (self-duality).

formed in this way is the same as the original lattice, we say that the lattice is *self-dual.*

Consider a closed polygon that appears in the evaluation of the partition function of Eq. (3.5). We put plus spins on the lattice sites inside the polygon and minus spins outside, and then connect opposite-spin pairs thus formed. Then, each side of the closed polygon is crossed by one and only one opposite-

spin connection. The opposite-spin connections are in the *dual* lattice when the polygon is in the *original* lattice. Therefore, we conclude that the partition function can be evaluated based on Eq. (3.5) by counting closed polygons on the original lattice, or based on Eq. (3.7) by counting opposite-spin pairs on the dual lattice.

These two methods must be equivalent. However, at high temperatures, the number of closed polygons is small, while at low temperatures, opposite-spin pairs are scarce (for ferromagnetic cases). Therefore, there can be practical differences between the two. If there exists just one critical point, the differences will disappear there.

In order to establish one-to-one correspondence between the partition functions evaluated by these two methods, let us consider a square lattice. Since this lattice is self-dual, its treatment is simple. We try to transform Eq. (3.5) into Eq. (3.7) or vice versa. For this purpose, let us define a function K^* of K by

$$\tanh K^* = e^{-2K} \tag{3.8}$$

remembering that x in Eq. (3.5) represents $\tanh K$.

Using Eq. (3.8) in Eq. (3.7), we find

$$\begin{aligned} Ze^{-SK} &= 2\left(1 + \sum_r \omega_r e^{-2Kr}\right) \\ &= 2\left[1 + \sum_r \Omega_r(\tanh K^*)^r\right], \end{aligned} \tag{3.9}$$

where the symbol $\Omega_r = \omega_r$ has been introduced so that the right-hand series assumes the form of the series in Eq. (3.5). We can now establish a formal relation between Z at K and Z^* at K^*, Z^* being the partition expressed by Eq. (3.5). By comparing Eqs. (3.9) and (3.5) we find

$$Ze^{-SK} = Z^*(K^*)[2^{N^*-1}(\cosh K^*)^S]^{-1}, \tag{3.10}$$

where Z must be considered to be a function of K because of the first right-hand side of Eq. (3.9).

The function $K^* = K^*(K)$ defined in Eq. (3.8) satisfies the expression

$$\sinh 2K \sinh 2K^* = 1, \tag{3.11}$$

which suggests correspondence relations such as

$$\begin{aligned} \text{large} \quad K &\rightleftarrows \text{small} \quad K^*, \\ \text{low-temp.} \quad Z &\rightleftarrows \text{high-temp.} \quad Z^*. \end{aligned} \tag{3.12}$$

The function $K^*(K)$ satisfies

$$\cosh 2K \tanh 2K^* = \cosh 2K^* \tanh 2K = 1. \tag{3.13}$$

With use of these relations, Eq. (3.10) can be brought into a more symmetric form:

$$\frac{Z(T)}{2^{N/2}(\cosh 2K)^{S/2}} = \frac{Z^*(T^*)}{2^{N^*/2}(\cosh 2K^*)^{S/2}}, \tag{3.14}$$

where T^* is the temperature corresponding to $K^* = J/kT^*$ and where

$$N + N^* = S + 2 \quad \text{and} \quad S^* = S.$$

From the symmetry of Eq. (3.14), which is the result of the self-duality of the square lattice, and from the correspondence relations (3.11) and (3.12), a singularity is determined by

$$\sinh 2K_c^* = 1, \tag{3.15}$$

if it exists. The subscript "c" stands for "critical." This relation gives

$$e^{-2K_c} = \sqrt{2} - 1. \tag{3.16}$$

Equation (3.16) is the exact relation which determines the critical temperature for a rectangular lattice. It gives

$$K_c \sim 0.4407, \tag{3.17}$$

where $K_c = J/2kT_c$. Thus, if the exchange integral J is known, T_c is determined without any adjustable parameter.

However, it should be remembered that the above theory did not prove the existence of a critical temperature. It merely determines T_c if it exists. The next section shall outline an exact theory for a two-dimensional rectangular lattice.

8.4.* Rigorous Theory of a Two-Dimensional Rectangular Lattice

The Onsager theory (*12*) was constructed with abstract mathematics, based on deep physical insight. Since then, several attempts have been made to simplify the theory. We shall outline in several steps one of the latest approaches to the problem.

8.4.1. Terminal Couplings

Let us consider a square lattice on a torus, so that edge effects can be neglected. Let $N = M \times M$ be the total number of lattice sites. According to Eq. (3.5), the partition function may be expressed in the form

$$Z = 2^N(\cosh K)^{2N}\left(1 + \sum_n \Omega_n x^n\right), \tag{4.1}$$

where $x = \tanh K$ and Ω_n is the number of closed polygons of n sides, and where $S = 2N$ has been used.

Each lattice site can have zero, two, or four bonds. No three-bond case appears because odd powers of σ do not contribute to Z. The bonds are in four directions, which may be called E, N, W, and S. We introduce four terminals E_j, N_j, W_j, and S_j for a site $j = 1, \ldots, N$. The closed polygons entering in Eq. (4.1) are now drawn such that each bond ends at the closest terminals of the nearest-neighboring sites (Fig. 8.3). For instance, the terminal E_j may have a bond with W_{j+1}, and N_j a bond with S_J. Here, J is the northern nearest-neighbor site to the site j. Thus, we obtain such polygons as shown by solid lines in Fig. 8.3. The dotted lines at each vertex close the polygon, and we shall explain them later.

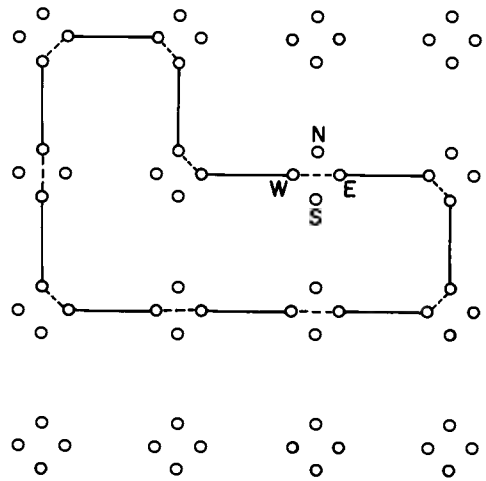


Fig. 8.3. A polygon in a terminaled lattice. Each lattice site has four terminals. A pair of the four terminals at a lattice site with solid bonds is connected by a broken bond.

The numbering of the terminals shall be made in the order E_j, N_j, W_j, and S_j following $4j-3$, $4j-2$, $4j-1$, and $4j$. Corresponding to the couplings between these terminals, we introduce the coupling factors:

1. For bondings,

$$a(E_j, W_{j+1}) = a(N_j, S_j) = -x; \qquad x = \tanh K. \tag{4.2}$$

2. For terminal closings,

$$a(E_j, N_j) = a(W_j, S_j) = 1; \qquad a(N_j, W_j) = a(E_j, S_j) = 1; \qquad a(E_j, W_j) = a(N_j, S_j) = -1. \tag{4.3}$$

The term "terminal closings" means connections between the terminals of the same lattice point. The four terminals can be closed by two bonds in three different modes as shown in Fig. 8.4. The mode (c) corresponds to the last of Eqs. (4.3). We shall call such a mode a *cross*.

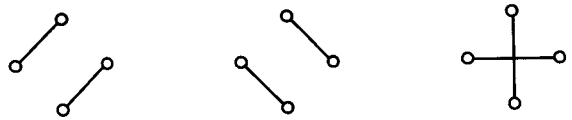


Fig. 8.4. The three modes of terminal closing.

The purpose of introducing these terminals is to transform the counting statistics of closed polygons into the coupling stastistics of these terminals. At first sight, the latter look more complicated because of the three extra processes of closing terminals. However, thanks to the minus sign given to a cross coupling, the problem is greatly simplified. In fact, the expansion (4.1) is reduced to

$$Z = 2^N(\cosh K)^{2N}\left[1 + \sum_n \sum_c (-)^c g(c, n)x^n\right]. \tag{4.4}$$

and further to a determinant. In this expression, $g(c, n)$ is the number of different configurations for n bonds with c crosses. In order to relate this expression to a determinant, we introduce a second step as follows.

8.4.2. Pfaffian Representation

Using the a's defined by Eqs. (4.2) and (4.3) as elements, one can construct the following pattern:

	E_1	N_1	W_1	S_1	E_2	N_2	W_2	S_2	E_3	N_3	W_3	S_3
E_1	1	−1	1		0	0	$-x$	0		zeros		
N_1		1	−1		0	0	0	0				
W_1			1		0	0	0	0				
S_1					0	0	0	0				
E_2					1	−1	1		0	0	$-x$	0
N_2						1	−1		0	0	0	0
W_2							1		0	0	0	0
S_2									0	0	0	0

This type of pattern is what is known in mathematics as a *Pfaffian*. The value of a Pfaffian is defined by the sum of products of the elements a_{ij} in the following form:

$$\{|A\} = \sum (-)^P a_{1i}a_{jk}\cdots a_{lm}, \tag{4.5}$$

Here, the symbol $\{|A\}$ represents the value, P represents the number of permutations, and m is the order of the Pfaffian. In the case of a square lattice of $M \times M$ elements, $m = 4M^2 = 4N$. Each element a_{ij} is placed in accordance with $i < j \leqslant m$. Under this condition, the right-hand side of Eq. (4.5) is obtained by permuting all the integers 1, 2, ..., m, giving odd permutations a minus sign and even permutations a plus sign. Thus, each product has $m/2$ factors in which all subscripts are different from each other.

A Pfaffian is related to an antisymmetric determinant of the elements of the following type:

$$a_{rr} = 0, \qquad a_{ij} = -a_{ji}. \tag{4.6}$$

Then, one can show

$$\{|\mathbf{A}\}^2 = |\mathbf{A}|, \tag{4.7}$$

where the right-hand side represents $\det|a_{rs}|$.

Let us omit the proof of this equality and see its validity for two special cases. For $m = 2$, we have

$$|\mathbf{A}| = \begin{vmatrix} 0 & a_{12} \\ a_{21} & 0 \end{vmatrix} = (a_{12})^2.$$

If $m = 4$, we find

$$\begin{aligned} |\mathbf{A}| &= \sum (-)^P a_{1i}a_{jk}a_{rs}a_{t4} \\ &= (a_{12}a_{34} - a_{13}a_{24} + a_{14}a_{23})^2. \end{aligned}$$

An antisymmetric determinant of zero diagonal elements vanishes for odd orders. In our case, m cannot be odd since the total number of terminals is even.

In the Pfaffian representation, the partition function of Eq. (4.4) becomes

$$Z = 2^N(\cosh K)^{2N}\{|\mathbf{A}\}. \tag{4.8}$$

In the remainder of this section, we shall give an outline of the proof of Eq. (4.8).

A lattice site is either bonded or nonbonded. If a lattice site is not bonded, its four terminals are *closed* in accordance with Eqs. (4.3). Note that in the Pfaffian expansion an even permutation of the subscripts in the product

$$a(E_j, N_j)a(W_j, S_j)$$

yields

$$a(N_j, W_j)a(E_j, S_j).$$

Both products occur in the expansion and have the same sign. However, the product

$$a(E_j, W_j)a(N_j, S_j)$$

arises from an odd permutation. Since all these products are equal to 1 by Eq. (4.3), the last crossing mode compensates for the extra weight factor of the former two couplings taken together. Therefore, any nonbonded lattice site can be correctly taken into account in the Pfaffian expansion by the terminal closing processes illustrated in Fig. 8.4.

A lattice site may be a vertex of a closed polygon. In Fig. 8.3, we have illustrated a polygon by solid lines between the terminals. This polygon is not closed, because at each lattice site none of the internal connections between its terminals has yet been made. One can always close a polygon by introducing intrasite terminal couplings such as shown by dotted lines in Fig. 8.3. The result is a new polygon with all the corners "cut" by dotted lines.

In introducing intrasite terminal couplings, we follow the order

$$E_j < N_j < W_j < S_j.$$

In other words, E_j may be coupled with N_j, W_j, or S_j, and N_j may be connected with W_j or S_j, while W_j may be given a bond with S_j. Therefore, the subscripts of any product of the elements a's corresponding to the dotted lines of the new polygon must form a sequence which can be rearranged in an ascending order by even permutations. Moreover, the number of crosses is always even. Thus, we conclude that the contribution from such a product to the Pfaffian is 1.

The product of the a's corresponding to all the solid-line and dotted-line couplings of the new polygon forms a cycle in which each terminal appears twice, once in a dotted-line coupling and once in a solid-line coupling. Thus, that part of the product which corresponds only to the solid-line couplings can be obtained by permutations of the suffices of the dotted-line couplings. As can be seen, these permutations are even. In addition, in a solid-line polygon, a coupling running from N to S (E to W) is always accompanied by another coupling running from S to N (W to E). Therefore, we conclude that the contribution from the product of a's of solid-line couplings to the Pfaffian is x^n, where n is the number of couplings. This result is in agreement with the fact that a polygon in a square lattice has an even number of sides and that each side (coupling) contributes $-x$ to the Pfaffian.

Finally, the product of the a's appearing by terminal closing processes (by dotted lines) at each lattice site contributes 1 to the Pfaffian. Since each term in the Pfaffian expansion exhausts all possible terminal couplings (spin interactions), we arrive at the partition function of Eq. (4.8).

8.4.3. Diagonalization

The antisymmetric matrix **A** associated with the Pfaffian shown on page 224 may be diagonalized. The size of **A** is $4M \times 4M$, and the eigenvectors are $4M$-dimensional. The components of the eigenvectors belonging to the eigenvalue $\lambda^{(\alpha)}$ may be represented as

$$\psi^{(\alpha)}(E_i), \quad \psi^{(\alpha)}(N_i), \quad \psi^{(\alpha)}(W_i), \quad \psi^{(\alpha)}(S_i), \qquad i = 1, 2, \ldots, N.$$

The eigenvalue problem is expressed by

$$\sum_i [a(Q_j, E_i)\psi^{(\alpha)}(E_i) + a(Q_j, N_i)\psi^{(\alpha)}(N_i) + a(Q_j, W_i)\psi^{(\alpha)}(W_i) + a(Q_j, S_i)\psi^{(\alpha)}(S_i)] = \lambda^{(\alpha)}\psi^{(\alpha)}(Q_j), \tag{4.9}$$

where Q stands for E, N, W, and S. Because of the form of **A**, it is possible to rewrite Eq. (4.9) in the following way:

$$\mathbf{U}\boldsymbol{\psi}_r^{(\alpha)} - \mathbf{X}\boldsymbol{\psi}_{r+1}^{(\alpha)} + \mathbf{X}'\boldsymbol{\psi}_{r-1}^{(\alpha)} - \mathbf{Y}\boldsymbol{\psi}_{r+4N}^{(\alpha)} + \mathbf{Y}'\boldsymbol{\psi}_{r-4N}^{(\alpha)} = \lambda^{(\alpha)}\boldsymbol{\psi}_r^{(\alpha)}, \qquad r = 1, 2, \ldots, N. \tag{4.10}$$

Here, the $\boldsymbol{\psi}^{(\alpha)}$ are four-dimensional, and

$$\mathbf{U} = \begin{bmatrix} 0 & 1 & -1 & 1 \\ -1 & 0 & 1 & -1 \\ 1 & -1 & 0 & 1 \\ -1 & 1 & -1 & 0 \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} 0 & 0 & x & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathbf{Y} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & x \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

$\mathbf{X}'$ and $\mathbf{Y}'$ are transposed matrices of **X** and **Y**, respectively. The vectors $\boldsymbol{\psi}_r^{(\alpha)}$ have four components. For negative subscripts r, $\boldsymbol{\psi}_r^{(\alpha)}$ is given the value $-\boldsymbol{\psi}_{N+r}^{(\alpha)}$:

$$\boldsymbol{\psi}_r^{(\alpha)} = -\boldsymbol{\psi}_{N+r}^{(\alpha)}. \tag{4.11}$$

Due to this property, one finds the proportionality

$$\boldsymbol{\psi}_r^{(\alpha)} \propto e^{r(2k-1)\pi i/N}. \tag{4.12}$$

The eigenvalue equation (4.10) is reduced to a simple form. The eigenstates are specified by the integer k, which runs from 1 to N. Once the eigenvalues are found, we can use

$$|\mathbf{A}| = \prod_k^N \lambda_k, \qquad \log|\mathbf{A}| = \sum_k^N \log \lambda_k. \tag{4.13}$$

In view of the form of the exponential function in (4.12), we introduce

$$k = (j-1)M + i,$$

where $1 \leqslant j \leqslant M$ and $1 \leqslant i \leqslant M$. Then,

$$\frac{2k-1}{N} = \frac{2(j-1)}{M} + \frac{2i-1}{M^2}.$$

In the limit $M \to \infty$, one can introduce continuous variables

$$\phi = \frac{(2i-1)\pi i}{M^2}, \qquad \theta = \frac{2(j-1)\pi i}{M} + \frac{\phi}{M}, \tag{4.14}$$

such that

$$\sum_k \to \frac{M}{2\pi}\int_0^{2\pi} d\theta \frac{M}{2\pi}\int_0^{2\pi} d\phi.$$

8.4.4. Rigorous Solution

Using the procedures described above, we finally arrive at the results obtained first by Onsager. Equation (4.13) for the two-dimensional square Ising lattice assumes the form

$$\log|\mathbf{A}| = \frac{N}{4\pi^2}\int_0^{2\pi}\int_0^{2\pi} \log[(1+x^2)^2 - 2x(\cos\theta)(1-x^2) - 2x(\cos\phi)(1-x^2)]\, d\theta\, d\phi, \tag{4.15}$$

where $x = \tanh K$ as before. Thus,

$$\frac{1}{N}\log Z = \log(2\cosh 2K) + \frac{1}{2\pi^2}\iint_0^{\pi} \log[1 - 4\kappa\cos\theta\cos\phi]\, d\theta\, d\phi, \tag{4.16}$$

where

$$\kappa = \frac{2\sinh 2K}{\cosh^2 2K}. \tag{4.17}$$

The internal energy is

$$U = -\frac{N}{2}J(\coth 2K)\left[1 \pm \frac{2}{\pi}(1-\kappa^2)^{1/2}J(\kappa)\right] \tag{4.18}$$

above (+) and below (−) T_c, where

$$J(\kappa) = \int_0^{\pi/2} \frac{d\phi}{(1-\kappa^2\sin^2\phi)^{1/2}}. \tag{4.19}$$

The above results show that a phase transition indeed exists at a temperature determined by Eq. (3.14). In the vicinity of T_c,

$$J(\kappa) \propto \log|T - T_c|, \qquad T < T_c. \tag{4.20}$$

Therefore, the energy is finite:

$$U \propto -(T - T_c)\log|T - T_c|, \tag{4.21}$$

and the heat capacity diverges logarithmically:

$$C_V \propto \log|T - T_c|. \tag{4.22}$$

This logarithmic singularity agrees with the recent experimental data on argon and other systems at the critical temperature.

The spontaneous magnetization is given by

$$M = \left[1 - \left(\sinh\frac{2J_1}{kT}\sinh\frac{2J_2}{kT}\right)^{-2}\right]^{1/8}, \tag{4.23}$$

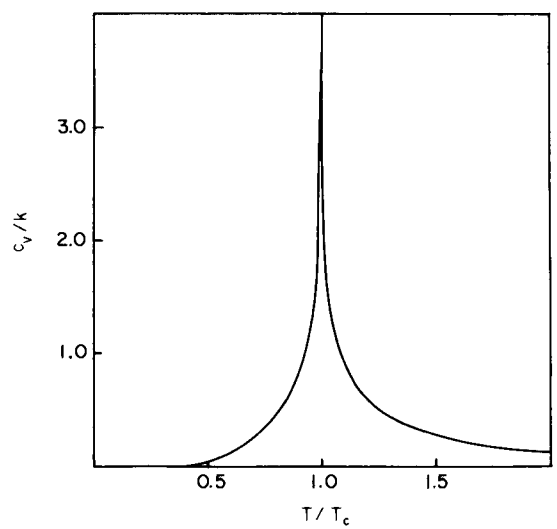


Fig. 8.5. The heat capacity curve of an Ising lattice based on Onsager's rigorous solution (two-dimensional isotropic square lattice).

if $J_1 \neq J_2$, that is, if the lattice is anisotropic with respect to spin couplings. This result was reported by Onsager at a conference held in 1948 at Cornell University, but the detailed proof was given by C. N. Yang in 1952 (*13*). Near T_c, the magnetization varies as

$$M \propto (T_c - T)^{1/8}. \tag{4.24}$$

Furthermore, computational and other analyses have shown that the susceptibility χ near T_c for an Ising lattice varies as [cf. Eq. (5.11)]

$$\chi \propto (T - T_c)^{-7/4}. \tag{4.25}$$

Figure 8.5 illustrates the Onsager result for the specific heat at constant volume. (For other cases, see General References.) Onsager obtained similar curves for the same quantity for isotropic and anisotropic exchange coupling constants. In the latter case, that is when the crystal has different exchange integrals for the two directions in a rectangular lattice, the heat capacity curve shows asymmetry below and above T_c.

8.5. Spin Correlation Functions

The correlation functions of spins are important because they show the order–disorder change more clearly than the partition functions. As in the case of fluids, there are many correlation functions, but the spin pair correlation function is most important both for equilibrium and nonequilibrium cases.

The spin variable σ takes on values 1 and -1 with equal probability since the configurational energy of the system is a quadratic function. However, if it is known that the jth lattice site has a spin σ_j, the probability of finding the spin σ_k at the kth site is no longer equal to $\frac{1}{2}$. This is simply because there exist correlations between the spins, which may be expressed by the quantity $\langle \sigma_i \sigma_j \rangle$, in comparison with $\langle \sigma_i \rangle \langle \sigma_j \rangle$. For an isotropic lattice, $\langle \sigma_i \sigma_j \rangle$ must be a function of the distance $|i - j|$. It is defined by (*14*, *15*)

$$\langle \sigma_1 \sigma_{1+r} \rangle = \frac{(\cosh K)^{2N}}{Z} \sum_{\sigma} \sigma_1 \sigma_{1+r} \prod_{(ij)} (1 + x\sigma_i \sigma_j). \tag{5.1}$$

The evaluation of the spin correlation function has been made by Montroll *et al.* (*14*) and others (*15*). Let us rewrite the right-hand side of Eq. (5.1) by making use of identities such as

$$\sigma_\alpha \sigma_i (1 + x\sigma_\alpha \sigma_l) = x(1 + x^{-1} \sigma_\alpha \sigma_i), \tag{5.2}$$

which arise because $\sigma_i^2 = 1$. We obtain

$$\langle \sigma_1 \sigma_{1+r} \rangle = \frac{(\cosh K)^{2N}}{Z x^N} \sum \prod_{(1k)} (1 + x^{-1} \sigma_1 \sigma_k) \prod_{(1+r\ l)} (1 + x^{-1} \sigma_{1+r} \sigma_l) \times \prod_{(ij)}{}' (1 + x\sigma_i \sigma_j), \tag{5.3}$$

where the last product is over all the nearest-neighbor spin pairs not included in the first two products which are for those of the two selected spins.

Therefore, the correlation function is obtained by a method similar to that for the partition function, with the only modification that the nearest neighbors of the two spins σ_1, σ_{1+r} must be given a factor x^{-1} instead of the regular factor x.

Our considerations in the previous section give immediately

$$|\mathbf{A}| x^{-2r} \langle \sigma_1 \sigma_{1+r} \rangle^2 = |\mathbf{A} + \Delta|, \tag{5.4}$$

where $|\mathbf{A}|$ is the determinant of the antisymmetric matrix (a_{rs}) and is given by Eq. (4.7), and $|\mathbf{A} + \Delta|$ is a new determinant constructed in accordance with the above replacement of x by x^{-1}, Δ representing the change due to the replacement. Equation (5.4) yields the formula

$$\langle \sigma_1 \sigma_{1+r} \rangle^2 = x^{2r} |\mathbf{1} + \mathbf{A}^{-1} \Delta|, \tag{5.5}$$

where $\mathbf{1}$ is a unit matrix of order $4N$ and $\mathbf{A}^{-1}$ is the inverse of the matrix $\mathbf{A} = (a_{rs})$.

The evaluation of the correlation function is now reduced to that of the inverse matrix. In general, when the eigenvalues λ_k and the eigenfunctions ψ_k of the original matrix $\mathbf{A}$ are known, the elements of the inverse matrix are obtained by

$$(\mathbf{A}^{-1})_{rs} = \sum_k \frac{1}{\lambda_k} \psi_k^*(r) \psi_k(s). \tag{5.6}$$

Using this formula in Eq. (5.5), one can evaluate the correlation function. Unfortunately, the evaluation is not very simple except for some particular cases. We only mention that the spontaneous magnetization M is related to the correlation function by

$$M^2 = \lim_{r \to \infty} \langle \sigma_1 \sigma_{r+1} \rangle. \tag{5.7}$$

It has been found that, for large r ($\varepsilon r \gg 1$),

$$\langle \sigma_1 \sigma_{1+r} \rangle = (2\varepsilon \cosh 2K)^{1/4} \frac{e^{-2\varepsilon r}}{(8\pi \varepsilon^2 r^2)} \qquad (T \lesssim T_c),$$

$$\cdot (2\varepsilon \cosh 2K)^{1/4} \frac{e^{-\varepsilon r}}{(2\pi \varepsilon r)^{1/2}} \qquad (T \gtrsim T_c), \tag{5.8}$$

where

$$\varepsilon = \frac{2|K^* - K|}{\sinh 2K}, \tag{5.9}$$

and K^* has been defined by Eq. (3.8), that is, $\tanh K^* = e^{-2K}$.

Because of the relation $\sinh 2K_c^* = \sinh 2K_c = 1$, ε is proportional to $(T_c - T)$. Equation (5.8) is valid for $\varepsilon r \gg 1$. Thus, as $T \to T_c$, i.e., $\varepsilon \to 0$, the correlation reaches longer and longer distances. Actually, one should take the limit $r \to \infty$ in approaching T_c. Then, as Eq. (5.7) states,

$$\lim_{r\to\infty} \langle \sigma_1 \sigma_{1+r} \rangle = M^2. \tag{5.10}$$

In the immediate vicinity of T_c, if $\varepsilon r \to 0$, the correlation between a diagonal pair varies as $r^{-1/4}$.

The results for small εr suggest $\chi \propto \varepsilon^{-7/4}$, because the zero-field susceptibility is given by

$$\begin{aligned} \chi_0 &= \beta\sigma^2 \sum_{(ij)} \langle \sigma_i \sigma_j \rangle \\ &\propto 2\pi \int \frac{e^{-2\varepsilon r} r\, dr}{r^{1/4}} \sim \varepsilon^{-7/4}, \end{aligned} \tag{5.11}$$

where in the second expression the summation over spin pairs is replaced by an integration. Also, it has been shown that, at T_c,

$$\lim_{r\to\infty} \langle \sigma_1 \sigma_{1+r} \rangle = 0, \tag{5.12}$$

as it should.

As we shall see later, time-dependent correlation functions characterize transport coefficients. For instance, the dynamic susceptibility is given by

$$\chi(\omega) = \beta \int_0^\infty \langle M(0)M(t) \rangle e^{-i\omega t}\, dt, \tag{5.13}$$

where $M(0)$ is the magnetization at $t = 0$ when the system is in equilibrium and $M(t)$ is that for a later time. The evaluation of the dynamic susceptibility may be attempted by assuming a master equation for the time dependence of the spin probability function.

8.6. Lattice Gas

In an Ising lattice, plus and minus spins may be considered to represent vacancies and gas atoms. The result is a lattice model for a gas, and the Ising lattice statistics can be used to find lattice gas properties.

The formal analogy between a lattice gas and a ferromagnet has been presented by Lee and Yang (*16*). Let us introduce the notation $[\uparrow]$ to represent the number of plus spins (vacancies in this case). The total number of lattice sites is denoted by

$$N = [\uparrow] + [\downarrow]. \tag{6.1}$$

The total energy of an Ising ferromagnet is

$$E_I = H([\downarrow] - [\uparrow]) - [\uparrow\downarrow]\varepsilon. \tag{6.2}$$

Therefore, the free energy per minus spin, f, is given by

$$e^{-Nf/kT} = \sum e^{-E_I/kT}, \tag{6.3}$$

where the summation is over all spin configurations. From the free energy, one obtains the intensity of magnetization per lattice site as follows:

$$I = -\frac{\partial f}{\partial H} = N^{-1} \langle \{[\uparrow] - [\downarrow]\} \rangle. \tag{6.4}$$

In the case of a lattice gas, the quantity

$$v = \frac{N}{\langle [\downarrow] \rangle} \tag{6.5}$$

represents the specific volume. Equation (6.4) leads to the relation

$$v = \frac{2}{1 - I}. \tag{6.6}$$

The energy of a lattice gas in the nearest-neighbor approximation can be expressed by

$$E_g = -[\downarrow\downarrow]\phi. \tag{6.7}$$

Thus, the grand partition function is

$$e^{pV/kT} = \sum e^{[\downarrow\downarrow]\phi/kT} z^{[\downarrow]}. \tag{6.8}$$

Let us try to relate Eq. (6.3) to Eq. (6.8). For this purpose, we rewrite the energy of a ferromagnet as follows:

$$-E_I = NH - 2H[\downarrow] - [\uparrow\downarrow]\varepsilon. \tag{6.9}$$

If the lattice is rectangular and two dimensional, each spin has four nearest neighbors, and we find

$$2[\downarrow\downarrow] + [\uparrow\downarrow] = 4[\downarrow]. \tag{6.10}$$

Rewriting Eq. (6.9) and using Eq. (6.3), we obtain

$$e^{-N(f+H)/kT} = \sum e^{2\varepsilon[\downarrow\downarrow]/kT} z^{[\downarrow]}, \tag{6.11}$$

where the summation goes over all the numbers of down spins, and where

$$z = e^{(-2H-4\varepsilon)/kT}. \tag{6.12}$$

Equation (6.11) coincides with Eq. (6.8) provided we set up the following correspondence relations:

$$\phi = 2\varepsilon, \qquad p = -f - H. \tag{6.13}$$

Therefore, we can transform the results on the Ising ferromagnet to those for a lattice gas. The relations are listed in Table 8.1.

TABLE 8.1

Relations between a Lattice Gas and an Ising Ferromagnet

Ising ferromagnet	Lattice gas
Total spin	Volume
$[\downarrow]$	Number of atoms
$2/(1 \cdot I)$	v
$-f \cdot H$	p

From Eqs. (6.4), (6.6), and (6.13) we find the relation

$$\begin{aligned} dp &= -df - dH = I\,dH - dH \\ &= -\frac{2}{v}\,dH. \end{aligned} \tag{6.14}$$

Thus,

$$H = -\frac{1}{2}\int v\,dp, \tag{6.15}$$

This relation shows how the p–v curves of a lattice gas are related to the magnetization curves.

Figure 8.6 illustrates the correspondence between the two systems. It is interesting to observe that the p–v curve looks like the I–H curve with the ordinate and abscissa interchanged. The magnetic field corresponds to $\frac{1}{2}$ of the shaded area $ABCD$ of (a). On the other hand, p of a lattice gas corresponds to the shaded area in (b) for a magnet, and the length ab is twice the density. For $T > T_c$, the Ising lattice has $I = 0$ at $H = 0$, and the corresponding pressure of the lattice gas is determined at $v = 2$. For $T < T_c$, the spontaneous magnetization at $H = 0$ of the Ising lattice corresponds to the volume at the condensation point of the lattice gas.

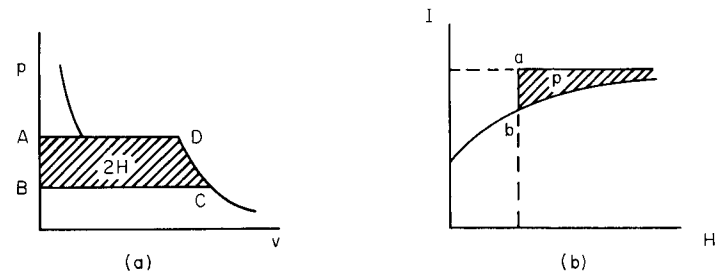


Fig. 8.6. The relation between the isotherm of a gas and the magnetization curve.

The results in this section can be combined with the exact Onsager solution to obtain the equation of state of a lattice gas. We arrive at the results (*16*)

$$\frac{p}{kT} = \log(1 + x^2) + \frac{1}{2\pi}\int_0^{\pi} \log\frac{1}{2}[1 + (1 - \kappa^2 \sin^2\phi)^{1/2}]\,d\phi; \tag{6.16}$$

$$v_g^{-1} = \frac{1}{2} - \frac{1}{2}\left[\frac{(1 + x^2)(1 - 6x^2 + x^4)^{1/2}}{(1 - x^2)^2}\right]^{1/4}, \tag{6.17}$$

where

$$\kappa = 4x(1 - x^2)(1 + x^2)^{-2} \tag{6.18}$$

and $x = e^{-\varepsilon/kT}$. In the equilibrium region, the specific volumes of the gas and liquid phases satisfy

$$v_g^{-1} + v_l^{-1} = 1. \tag{6.19}$$

In contrast, a *rectilinear relation* holds experimentally:

$$v_g^{-1} + v_l^{-1} = 2\left[1 + \frac{a(T_c - T)}{T_c}\right]v_c^{-1}, \tag{6.20}$$

where a is a constant and the subscript "c" means "critical."

For high temperatures, one can expand the pressure as follows:

$$\frac{p}{kT} = \sum_l b_l z^l, \tag{6.21}$$

$$b_1 = 1, \qquad b_2 = 2x^{-2} - \tfrac{5}{2}, \qquad b_3 = 6x^{-4} - 16x^{-2} + \tfrac{31}{3} \quad \ldots,$$

where z, the activity, is given by Eq. (6.12) for $H = 0$:

$$z = e^{-4\varepsilon/kT}. \tag{6.22}$$

Equation (6.21) is the expansion for small z. For large z, one finds the expansion in z^{-1}:

$$\frac{p}{kT} = \log\frac{z}{x^4} + \frac{x^8}{z} + \frac{1}{z^2}\left(2x^{14} - \frac{5}{2}x^{16}\right) + \cdots. \tag{6.23}$$

This expansion is derived with use of the symmetry of the I–H curve. A similar "cluster" expansion can be obtained for the density.

8.7. Distribution of Zeros of the Grand Partition Function

In Eq. (6.12), which relates the activity of a lattice gas with the magnetic field, the factor 4 in front of ε/kT came from the number of nearest neighbors. For a simple cubic lattice, this factor should be 6. In general, the activity relation can be written as

$$z = e^{(-2H-q\varepsilon)/kT}, \tag{7.1}$$

where q depends on lattice structure.

The free energy per particle of a lattice gas has been given by Eq. (6.11). Transferring the energy due to the magnetic field to the right-hand side, one obtains

$$e^{-Nf/kT} = \Xi e^{NH/kT}, \tag{7.2}$$

where Ξ, the grand partition function, is defined by the series of Eq. (6.11). Since the total number of spins is limited and since the Boltzmann factors are positive, the grand partition function Ξ is a polynomial in z with positive coefficients. Thus, one can write

$$\Xi = \sum_0^N Z_{[\downarrow]} z_1^{[\downarrow]} = \prod_k (z_1 - e^{i\theta_k}), \tag{7.3}$$

with $z_1 = e^{-2H/kT}$

In Eq. (7.3), $Z_{[\downarrow]}$ represents the partition function of the Ising lattice for zero magnetic field and for a given number $[\downarrow]$ of down spins. Obviously,

$$Z_{[\downarrow]} = Z_{N-[\downarrow]}. \tag{7.4}$$

Since $Z_{[\downarrow]} > 0$, the zeros of Ξ must be complex. Only in the limit $N \to \infty$, can some real zeros appear. If a real zero appears, we expect a phase transition.

Following a mathematical theorem by Winter, Lee and Yang showed that all the zeros of the polynomial Ξ lie on the unit circle around the origin in the complex z_1 plane if the interaction ϕ between the lattice gas atoms is such that

$$\phi\begin{cases} = \infty, & \text{if two coincide,} \\ \leqslant 0, & \text{otherwise.} \end{cases}$$

This theorem makes no assumption on the range of interactions and the size and structure of the lattice. It is applicable to ferromagnetic spin interactions and is not limited to nearest-neighbor interactions (*16*).

The unit circle cuts the real, positive axis only at one point. Therefore, the lattice gas can undergo a phase transition only once. From Eqs. (7.2) and (7.3), one finds in the limit $N \to \infty$

$$-\frac{f}{kT} = \lim_{N\to\infty}\frac{1}{N}\log Z_N = \frac{H}{kT} + \int_0^{2\pi} g(\theta; z_1)\log(z_1 - e^{i\theta})\,d\theta$$
$$= \frac{H}{kT} + \int_0^{\pi} g(\theta; z_1)\log({z_1}^2 - 2z_1\cos\theta + 1)\,d\theta, \tag{7.5}$$

where $g(\theta)$ represents the distribution function of zeros of Ξ. The last form is obtained by combining the contribution from a zero with that from its conjugate complex. The distribution function $g(\theta; z_1)$ is normalized as follows:

$$\int_0^{\pi} g(\theta; z_1)\,d\theta = \tfrac{1}{2}, \tag{7.6}$$

because $g(\theta) = g(-\theta)$ and the domain of θ is from 0 to 2π.

Since $z_1 = e^{-2H/kT}$, the free energy per spin given by Eq. (7.5) may be expressed by

$$-\frac{f}{kT} = \int_0^{\pi} g(\theta; z_1)\log 2(\cosh h - \cos\theta)\,d\theta, \tag{7.7}$$

where $h = 2H/kT$.

The intensity of magnetization is obtained from Eq. (7.5) in the form

$$I = 1 - 4z_1\int_0^{\pi} g(\theta; z_1)\frac{z_1 - \cos\theta}{{z_1}^2 - 2z_1\cos\theta + 1}\,d\theta. \tag{7.8}$$

On the other hand, for a lattice gas, use of Table 8.1 and of Eq. (7.8) gives

$$\frac{p}{kT} = \int_0^{\pi} g(\theta; z_1)\log({z_1}^2 - 2z_1\cos\theta + 1)\,d\theta \tag{7.9}$$

and

$$n = 2z_1\int_0^{\pi} g(\theta; z_1)\frac{z_1 - \cos\theta}{{z_1}^2 - 2z_1\cos\theta + 1}\,d\theta. \tag{7.10}$$

Let us evaluate the integrals of Eqs. (7.7) and (7.8) approximately near the critical temperature for the case of small θ and h. Expanding functions such as $\cosh h$ and $\cos\theta$, we obtain

$$-\frac{f}{kT} \approx \int_0^{\pi} g(\theta; z_1)\log(\theta^2 + h^2)\,d\theta, \tag{7.11}$$

and

$$I \propto 4h\int_0^{\pi}\frac{g(\theta; z_1)}{\theta^2 + h^2}\,d\theta. \tag{7.12}$$

In the limit $h \to 0$,

$$I_0 \propto 2\pi g(0; z_1), \tag{7.13}$$

I_0 should be zero for $T > T_c$.

Equation (7.12) gives the susceptibility for zero magnetic field as follows:

$$\chi_0^+ \propto \frac{8}{kT}\int_0^\pi \frac{g(\theta; z_1)}{\theta^2}\,d\theta, \qquad T > T_c, \tag{7.14}$$

and for $T < T_c$,

$$\chi_0^- = \lim_{H\to 0}\frac{I(H) - I_0}{H}$$
$$\propto \frac{8}{kT}\int_0^\pi \frac{g(\theta; z_1) - g(0; z_1)}{\theta^2}\,d\theta, \qquad T < T_c. \tag{7.15}$$

The heat capacity is determined by

$$C \propto \frac{d^2}{dT^2}\int_0^\pi g(\theta; z_1)\log\theta\,d\theta. \tag{7.16}$$

Further calculations require an explicit form for $g(\theta; z_1)$. However, the determination of $g(\theta; z_1)$ is equivalent to solving the Ising lattice problem. In the presence of a magnetic field, the one-dimensional case has been solved exactly. The zeros of Ξ are determined by

$$\cos\theta_j = -x^2 + (1 - x^2)\cos\frac{\pi(2j-1)}{N},$$

where $x = e^{-\epsilon/kT}$, and j stands for all integers equal to or less than $(N+1)/2$. As $N \to \infty$, the zeros are distributed on an arc of the unit circle lying to the left of the points

$$z_1 = (1 - 2x^2) \pm i2x(1 - x^2)^{1/2}. \tag{7.17}$$

For all values of $x > 0$, no zeros move into the positive, real axis. Therefore, there is no critical point. The distribution function is

$$g(\theta; x)\begin{cases} = \dfrac{1}{2\pi}\dfrac{\sin\frac{1}{2}\theta}{(\sin^2\frac{1}{2}\theta - x^2)^{1/2}}, & \cos\theta < 1 - 2x^2, \\ = 0, & \cos\theta > 1 - 2x^2. \end{cases} \tag{7.18}$$

For three-dimensional cases, critical power laws, which will be discussed in the next section, give some clue to finding a plausible form for $g(\theta; x)$. As experiments show, if the susceptibility and the magnetization vary as

$$\chi_0^+ \sim \tau^{-\gamma}, \qquad M \sim |\tau|^\beta, \tag{7.19}$$

for $\tau = (T - T_c)/T_c \to 0$, the distribution function $g(\theta; x)$ might have a form (*17*, *18a*, *18b*)

$$g(\theta; x) = \tau^\beta \phi\left(\frac{\theta}{\tau^{\beta+\gamma}}\right), \tag{7.20}$$

where ϕ is a new function of $\theta/\tau^{\beta+\gamma}$. The reader may want to try to find the consequences of Eq. (7.20).

8.8. Frequency Spectrum

To explain the heat capacity curves of solids, Debye introduced a frequency distribution function in a continuum model. This model is expected to be good at low temperatures, where oscillations of long wavelength are important. The Debye frequency distribution function is normalized to give the total degrees of freedom. Therefore, it is also supposed to be good at high temperatures, where the equipartition law holds.

In the intermediate-temperature region $T \sim \vartheta_D$, the Debye model is no longer very good, as can be seen from the fact that the characteristic temperature ϑ_D turns out to be temperature-dependent. Here, a more exact theory of lattice vibrations is needed. The first theory meeting this need was given by Born and von Karman in 1913 [see Maradudin *et al.* (*22*)].

The Born–von Karman theory appeared a year after the Debye theory. Since then, their model has been used widely. Refinements of their theory have been made by many investigators since around 1950.

A striking result of these investigations was the existence of certain singularities in the frequency distribution function $g(\omega)$ of the angular frequency. In 1953, Van Hove (*19*) showed that they are due to the lattice periodicity rather than to the forces between the atoms.

In the harmonic approximation, the displacement vector $\mathbf{u}(\mathbf{r})$ at the point $\mathbf{r} = (r_1, r_2, r_3)$ in a simple cubic lattice follows the equation of motion

$$m\ddot{\mathbf{u}} = \gamma_1 \Delta_x \mathbf{u}(\mathbf{r}) + \gamma_2 \Delta_y \mathbf{u}(\mathbf{r}) + \gamma_3 \Delta_z \mathbf{u}(\mathbf{r}), \tag{8.1}$$

where r_1, r_2, and r_3 are integers, the γ's are force constants, and where

$$\Delta_x \mathbf{u}(\mathbf{r}) = \mathbf{u}(r_1 + 1, r_2, r_3) - 2\mathbf{u}(r_1, r_2, r_3) + \mathbf{u}(r_1 - 1, r_2, r_3), \quad \text{etc.} \tag{8.2}$$

Let us assume a periodic boundary condition to avoid surface effects, and let us put

$$\mathbf{u}(\mathbf{r}) \propto \exp i\left(\frac{2\pi\mathbf{k}\cdot\mathbf{r}}{N} - \omega t\right), \tag{8.3}$$

where N is the number of the last atom and $\mathbf{k} = (k_1, k_2, k_3)$ is an integer which runs from 1 to N. Introducing Eq. (8.3) into Eq. (8.1), we find a dispersion relation

$$\omega^2 = 2\sum_{j=1}^{3} \frac{\gamma_j(1 - \cos\theta_j)}{m}, \tag{8.4}$$

where

$$\theta_j = \frac{2\pi k_j}{N}. \tag{8.5}$$

The frequency distribution function $g(\omega)\,d\omega$ is the fraction of possible vibrations of angular frequencies between ω and $\omega + d\omega$. The function $g(\omega)$ is normalized to one when integrated from zero to the maximum possible frequency. For the vibrations satisfying Eq. (8.4), the maximum frequency is

$$\omega_M{}^2 = \frac{4(\gamma_1 + \gamma_2 + \gamma_3)}{m}. \tag{8.6}$$

Because the dispersion relation (8.4) is in terms of ω^2, it is convenient to evaluate $G(\omega^2)\,d\omega^2$, which is the fraction of squared angular frequencies in the interval $(\omega^2, \omega^2 + d\omega^2)$. This function is related to $g(\omega)$ as follows:

$$g(\omega) = 2\omega G(\omega^2), \tag{8.7}$$

as can easily be seen. $G(\omega^2)$ is given by

$$\begin{aligned} G(\omega^2) &= \frac{1}{3N}\sum_{\mathbf{k}} \delta(\omega^2 - \omega^2(k)) \\ &= \frac{1}{3N}\sum_{\mathbf{k}} \frac{1}{2\pi}\int_{-\infty}^{\infty} \exp\{-i[\omega^2 - \omega^2(k)]x\}\,dx \\ &= \frac{1}{2\pi}\int \tilde{G}(x)\exp(-i\omega^2 x)\,dx, \end{aligned} \tag{8.8}$$

where

$$\tilde{G}(x) = \frac{1}{3N}\sum_{\mathbf{k}} \exp[i\omega^2(k)x]. \tag{8.9}$$

The Fourier transform $\tilde{G}(x)$ is evaluated by introducing Eq. (8.4) into Eq. (8.9). Replacing the summation over $\mathbf{k}$ by integration,

$$\begin{aligned} \tilde{G}(x) &= \prod_{j=1}^{3} \frac{1}{(2\pi)^3}\int_0^{2\pi} e^{2i\gamma_j x(1-\cos\theta_j)/m}\,d\theta_j \\ &= \prod_{j=1}^{3} e^{2i\gamma_j x/m} J_0\left(\frac{2i\gamma_j x}{m}\right). \end{aligned} \tag{8.10}$$

Thus,

$$G(\omega^2) = \frac{1}{2\pi}\int_{-\infty}^{\infty} (\exp -i\omega^2 x)\prod_{j=1}^{3}\left[\left(\exp\frac{2i\gamma_j x}{m}\right)J_0\left(\frac{2i\gamma_j x}{m}\right)\right] dx. \tag{8.11}$$

The integration is rather difficult to carry out analytically. A numerical result obtained by Montroll and co-workers (*20*, *22*) is illustrated in Fig. 8.7.

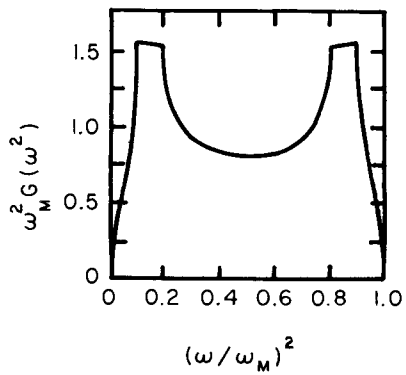


Fig. 8.7. The frequency spectrum of a three-dimensional simple cubic lattice.

For one- and two-dimensional cases, rigorous analytic results can be obtained.

The formula (8.11) is based on the approximation of nearest-neighbor interactions. However, it has been found that interactions between more

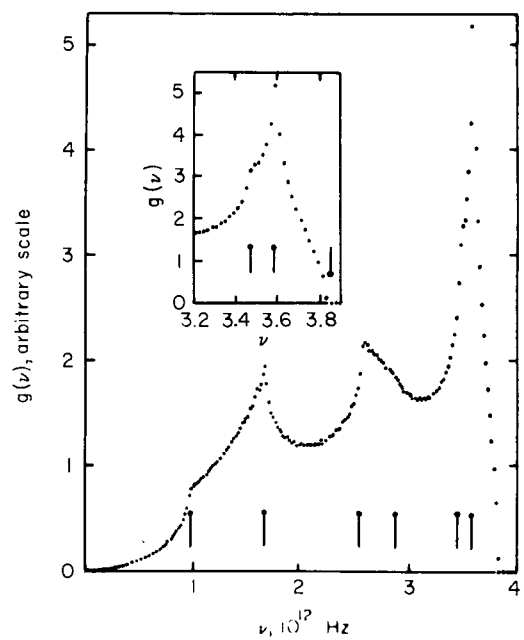


Fig. 8.8. The frequency distribution function of sodium. [After Dixon *et al.* (*23*).]

distant neighbors are necessary in order to provide a satisfactory description of experimental data. Also, electron–phonon interactions cannot be neglected in lead and many other substances.

The determination of the frequency distribution function has been performed for sodium, chlorine, aluminum, and other substances. Figure 8.8 gives the frequency distribution function of sodium determined by Dixon *et al.* (*23*) by neutron scattering. The locations of singularities are indicated by small bars. The details at the peak of the distribution function are shown in the inset. Sodium melts at 370°K, and a certain degree of anharmonicity exists even at 200°K. There are some other factors which make the experiment difficult, yet the frequency spectrum gives a very good specific heat curve.

Crystalline lithium fluoride and lithium hydride have simple NaCl-type face-centered cubic structures and have attracted much theoretical and experimental interest in recent years. The forces in these substances include the short-range overlap repulsions between neighbor ions, the Coulomb forces, and the electrostatic forces that arise because of distortions in the electronic configurations of ions due to phonons. Therefore, a simplification of the forces is desirable in order to avoid a large number of adjustable parameters.

Figure 8.9 shows the frequency distribution of normal modes in a single

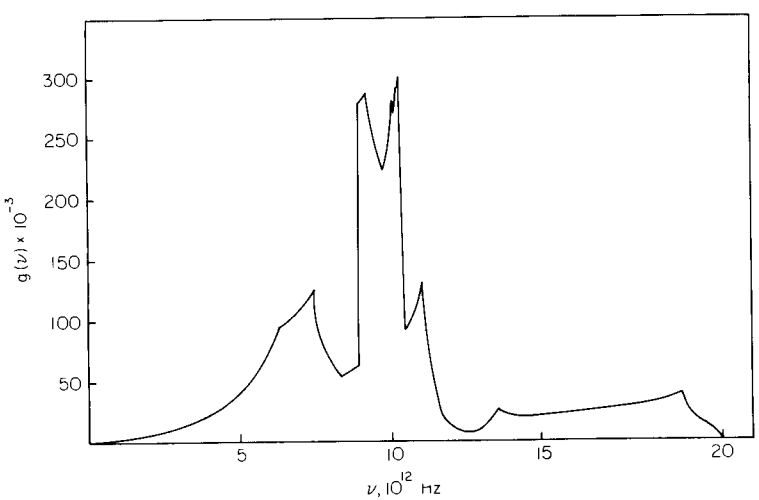


Fig. 8.9. The frequency distribution function of LiF. [After Dolling *et al.* (*24*).]

crystal of LiF determined from neutron scattering data by Dolling *et al.* (*24*). These authors (*25*, *26*) adopted a seven-parameter dipole approximation involving nearest-neighbor and next-nearest-neighbor $F^- - F^-$ short-range forces, a variable ionic charge, and the polarizability of the F^- ions. The forces between the Li^+ ions and their polarizability are both set equal to zero, in recognition of their smallness and tightly bound states.

The area under the frequency curve in Fig. 8.9 is normalized to one. The appearance of several singularities is noticed. From the curve, the heat capacity and the Debye temperature can be readily computed in the harmonic approximation. In Fig. 8.10, the Debye temperature thus obtained is plotted

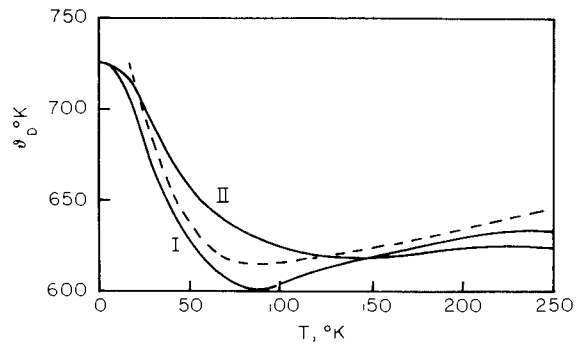


Fig. 8.10. The Debye characteristic temperature of LiF plotted against temperature.

against temperature. The two solid curves I and II are evaluated from the heat capacity and the entropy, respectively. The dashed curve shows the Debye temperature derived from the heat capacity data. Although anharmonic effects are neglected, the agreement is good except for temperature below 25°K.

The frequency distribution function yields all the thermodynamic functions of crystals. For instance, we have (*21*)

$$F = 3NkT \int_0^{\omega_M} \left[\log\left(2 \sinh \frac{\hbar\omega}{2kT}\right)\right] g(\omega)\, d\omega, \tag{8.12}$$

$$U = \frac{3N\hbar}{2} \int_0^{\omega_M} \omega \left(\cosh \frac{\hbar\omega}{2kT}\right) g(\omega)\, d\omega, \tag{8.13}$$

$$C_V = 3Nk \int_0^{\omega_M} \left(\frac{\hbar\omega}{2kT}\right)^2 \left(\cosh^2 \frac{\hbar\omega}{2kT}\right) g(\omega)\, d\omega. \tag{8.14}$$

Problems

1. Show that, for a one-dimensional lattice,

$$G(\omega^2) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-i\omega t} e^{2i\gamma} J_0(2i\gamma)\, dt$$

$$= \begin{cases} [\pi\omega(\omega_M{}^2 - \omega_2{}^2)^{1/2}]^{-1}, & \omega^2 < \omega_M{}^2, \\ 0, & \omega^2 > \omega_M{}^2, \end{cases} \tag{8.15}$$

where $\omega_M{}^2 = 4\gamma/m$.

2. Show that the sum of the type $\sum_j \Phi(\omega_j)$ over normal modes can be expressed by a contour integral

$$\langle\Phi\rangle = \sum_j \Phi(\omega_j) = \frac{1}{2\pi i}\oint \Phi(\omega)\frac{d}{d\omega}\left[\log\prod_j(\omega^2-\omega_j{}^2)\right]d\omega. \tag{8.16}$$

Thus, if

$$D(\omega) = \prod_j(\omega^2-\omega_j{}^2)$$

the effects of lattice defects on $\langle\Phi\rangle$ may be obtained from

$$\Delta\langle\Phi\rangle = \frac{1}{2\pi i}\oint \Phi(\omega)\frac{d\{\log[D(\omega)/D_0(\omega)]\}}{d\omega}d\omega, \tag{8.17}$$

where $D_0(\omega)$ represents $D(\omega)$ for the case of a perfect lattice. From this observation, write down the internal energy change ΔU due to defects.

8.9.* Lattice Green's Function

Let us consider a one-dimensional ring of N atoms. If the jth lattice site is a defect, such as an impurity or a vacancy, the equations of motion in the x direction are of the form

$$m\ddot{x}_j - \gamma(x_{j+1} - 2x_j + x_{j-1}) = D_j x_j, \qquad j = -\tfrac{1}{2}N+1, \ldots, \tfrac{1}{2}N, \tag{9.1}$$

$$x_{N+j} = x_j, \tag{9.2}$$

where x_j represents the displacement of the jth atom in the x direction from its equilibrium position and D_j is an operator which characterizes the lattice defects.

The condition $D_j = 0$ corresponds to a perfect lattice. The solution to Eqs. (9.1) can be expressed as a linear combination of the normal modes:

$$x_j = A_k a e^{i(t\omega_k + j\phi_k)}, \qquad \phi_k = \frac{2\pi k}{N}, \tag{9.3}$$

where the ω_k are the normal-mode frequencies given by

$$\omega_k = 2\left(\frac{\gamma}{m}\right)^{1/2}\sin\left(\frac{\phi_k}{2}\right)$$

$$= \omega_M \sin\left(\frac{\phi_k}{2}\right), \tag{9.4}$$

where $\omega_M = (4\gamma/m)^{1/2}$ is the largest frequency. Using Eq. (9.3), we can rewrite Eqs. (9.1) as

$$\begin{aligned} Lx_j &= m\omega^2 x_j + \gamma(x_{j+1} - 2x_j + x_{j-1}) \\ &= -D_j x_j, \end{aligned} \tag{9.5}$$

where L is an operator to produce the right-hand side of the first equation. The operator L satisfies

$$Lg(j) = \begin{cases} 1, & j = 0, \\ 0, & j \neq 0, \end{cases} \tag{9.6}$$

where g is the Green's function. It gives the solution of Eq. (9.5) in the form

$$x_j = -\sum_l g(j-l)D_l x_l, \tag{9.7}$$

because

$$Lx_j = -\sum_l Lg(j-l)D_l x_l = -D_j x_j. \tag{9.8}$$

The Green's function $g(j)$ can be found by writing the left-hand side of Eq. (9.6) explicitly and performing the sum:

$$\begin{aligned} \tilde{g}(\theta) &= \sum_j g(j)e^{i\theta j} \\ &= [m\omega^2 + 2\gamma(\cos\theta - 1)]^{-1}. \end{aligned} \tag{9.9}$$

$$\begin{aligned} g(j) &= \mathrm{Re}\frac{1}{2\pi}\int_0^{2\pi}\tilde{g}(\theta)e^{-i\theta j}\,d\theta \\ &= \frac{1}{\pi}\int_0^{\pi}\frac{\cos\theta j}{m\omega^2 + 2\gamma(\cos\theta - 1)}\,d\theta \\ &= \frac{(-)^j}{\gamma}\frac{e^{-|j|z}}{2\sinh z} \qquad (\omega > \omega_M), \end{aligned} \tag{9.10}$$

where $\cosh(z/2) = \omega/\omega_M$ and "Re" means the real part.

One can generalize the above consideration easily to the three-dimensional case. For a simple cubic lattice, we obtain

$$g(r) = \frac{1}{\pi^3}\int\frac{\cos r_1\theta_1 \cos r_2\theta_2 \cos r_3\theta_3}{m\omega^2 + 2\sum_{j=1}^3 \gamma_j(\cos\theta_j - 1)}\,d\theta_1 d\theta_2\,d\theta_3. \tag{9.11}$$

The function $g(r)$ of Eq. (9.11) is the Green's function for a simple cubic lattice. For other types of lattice, one gets similar expressions. For large values of (r_1, r_2, r_3) and for the case $\gamma_1 = \gamma_2 = \gamma_3$, g approaches

$$g(r) \sim \frac{\exp[-2\cdot 3^{1/2}(\omega/\omega_M)r]}{(-4\pi\gamma)r}, \qquad \omega > \omega_M = \left(\frac{12\gamma}{m}\right)^{1/2}, \tag{9.12}$$

where r_j are integers and

$$r^2 = r_1{}^2 + r_2{}^2 + r_3{}^2. \tag{9.13}$$

It is clear from Eq. (9.7) that the Green's function yields the solutions of the equations of motion. The effects of lattice defects and other factors can be found from the solutions (*20–22*).

8.10.* Spherical Model

In the so-called spherical model, the constraint $\sigma_j^2 = 1$ for each spin is replaced by a weaker one:

$$\sum_j \sigma_j^2 = N. \tag{10.1}$$

This constraint retains the correct average values $\langle \sigma_j^2 \rangle = 1$ and prevents large fluctuations of order N of $\langle \sigma_j^2 \rangle$ about this mean value.

This model, introduced by Berlin and Kac (*27*), may be considered as a limit of infinite spin dimensionality. It is only an approximation to the solution of the Ising problem, but has the advantage of mathematical tractability. At least at high temperatures it gives results similar to those from the Ising lattice. Moreover, the mathematical technique for the spherical model can be used for some other systems, such as electrolyte solutions.

The spherical model does not give a phase transition for two-dimensional lattices. However, it does for the three-dimensional cases.

The constraint of Eq. (10.1) may be expressed in terms of the Dirac delta function. Then, the partition function, which is normalized to unity, is given by

$$\begin{aligned} Z &= \frac{\Omega^{-1}}{2\pi i} \int \cdots \int \exp\left[K \sum_{(ij)}{}' \sigma_i \sigma_j\right] \delta\left(N - \sum_j \sigma_j^2\right) \prod_j d\sigma_j \\ &= \frac{\Omega^{-1}}{2\pi i} \int d\zeta (\exp N\zeta) \int d\sigma_1 \cdots d\sigma_N \exp\left(-\zeta \sum_j \sigma_j^2 + K \sum_{(ij)} \sigma_i \sigma_j\right), \end{aligned} \tag{10.2}$$

where $K = J/2kT$ and Ω is a normalization constant:

$$\Omega = \int \cdots \int d\sigma_1 \cdots d\sigma_N = \frac{2\pi^{N/2} N^{(N-1)/2}}{\Gamma(N/2)}. \tag{10.3}$$

The integral of the type given in Eq. (10.2) can be calculated easily by employing the formula

$$\int \cdots \int \exp\left(-\sum_{i,j=1}^{N} a_{ij} \sigma_i \sigma_j\right) \prod d\sigma_i = \pi^{N/2} (\det |a_{jk}|)^{-1/2}. \tag{10.4}$$

It is found that

$$\lim_{N \to \infty} \frac{1}{N} \log Z = -\frac{1}{2} - \frac{1}{2} \log 2K - \frac{1}{2} f(z_s) + K z_s, \tag{10.5}$$

where z_s represents the solution of

$$4K = f'(z) = \frac{1}{(2\pi)^3} \iiint \frac{d\omega_1\, d\omega_2\, d\omega_3}{z - \cos\omega_1 - \cos\omega_2 - \cos\omega_3} \tag{10.6}$$

for simple cubic lattices. $f(z)$ is the integral of the Green's function $f'(z)$.

Equation (10.6) has a real solution only when

$$T \geqslant T_c = 3.9568 \frac{J}{K}. \tag{10.7}$$

Singularities appear in the thermodynamic functions at T_c. For instance, the specific heat $(k/2)(1 + 2K^2\, dz_s/dK)$ increases continuously with decreasing temperature until $T = T_c$, at which it suddenly becomes a constant $k/2$ for all $T < T_c$. The spontaneous magnetization is proportional to $[1 - (T/T_c)]^{1/2}$ for $T < T_c$ as in the classical theory. The spin correlation function vanishes above T_c and varies as $[1 - (T/T_c)]$ below T_c.

The spherical model is useful since all these thermodynamic properties are evaluated analytically. Moreover, it can be applied to more complex cases. For instance, if a simple cubic lattice has defects which are distributed dilutely in the lattice with a concentration c, the Curie temperature is shifted by the amount (*28*)

$$\Delta T_c = 0.2308c(K'^2 - K^2) \frac{k T_c^2}{J K^2}. \tag{10.8}$$

That is, the Curie temperature varies linearly with the defect concentration. Here, K' is the interaction constant of the defect atoms.

8.11.* Heisenberg Model

The statistical mechanics of ferromagnetism based on the Heisenberg model has been analyzed for low temperatures in terms of a series expansion in powers of T (the spin-wave solution) and for higher temperatures in terms of a series in $1/T$ (*29*). Since these approaches are valid only at temperatures far from the critical region, efforts have been made to approach the critical region by Brout (see General References), Horwitz and Callen, and others (*30–32*) based on a cluster expansion and other methods.

The Hamiltonian of the Heisenberg model is given by

$$H = -\sum_{(ij)} J \mathbf{S}_i \cdot \mathbf{S}_j - H \sum_j S_{jz}, \tag{11.1}$$

where $\mathbf{S}_i$ is the total spin operator for the ith lattice site, S_{iz} being its component in the direction of the magnetic field H. The first summation is over the nearest-neighboring spin pairs and the second term represents the effect of

the external magnetic field. In so far as the Hamiltonian of Eq. (11.1) represents ferromagnetism, the spins must be lined up at 0°K. Indeed, one can prove that the Hamiltonian has a ground state in which all the spins have maximum alignment along, say, the z axis. Nevertheless, in one and two dimensions, the isotropic Heisenberg model does not show a spontaneous magnetization. One has to go to the three-dimensional case for a spontaneous magnetization (*33*).

8.11.1. Magnons (Spin-Wave Theory)

Imagine a small perturbation of the ground state of a ferromagnet which is caused by irregular orientation of one of the spins. The perturbed state cannot be an eigenstate of H because the spin deviation could be on some other lattice points. In other words, spin deviations keep moving in the lattice.

Spin deviations cost energy. Although individual excitations might be costly, the Heisenberg Hamiltonian allows much "cheaper" collective modes of excitation called *spin waves*.

Quantized spin waves are called *magnons*. Different from phonons, which have energies proportional to their momenta, the magnon energies in the lowest approximation may be represented by

$$\hbar\omega \sim 2SJq^2a^2,$$

where S is the maximum spin value at each site, q is the wave number, and a is the lattice constant. Because of the q^2 dependence, we can express the magnon energy in terms of effective mass m^* as follows:

$$\hbar\omega = \frac{\hbar^2 q^2}{2m^*}. \tag{11.2}$$

The spin waves can be superposed, and the magnons obey Bose–Einstein statistics. We then find the average magnon number

$$\begin{aligned} N &= \frac{V}{8\pi^3}\int_0^\infty \frac{4\pi q^2\,dq}{\exp(2SJa^2q^2/kT)-1} \\ &= \frac{N}{4\pi^2}\left(\frac{kT}{2SJ}\right)^{3/2}\int_0^\infty \frac{x^{1/2}\,dx}{e^x-1}, \end{aligned} \tag{11.3}$$

where $Na^3 = V$ has been used. The internal energy is given by

$$\begin{aligned} U &= \sum_q \hbar\omega\langle n\rangle \\ &= N\frac{(kT)^{5/2}}{4\pi^2}\left(\frac{1}{2SJ}\right)^{3/2}\int_0^\infty \frac{x^{3/2}\,dx}{e^x-1} \\ &= \frac{3N(kT)^{5/2}\zeta(5/2)}{2(8\pi SJ)^{3/2}}. \end{aligned} \tag{11.4}$$

Thus, the specific heat is proportional to $T^{3/2}$ (*34*):

$$\begin{aligned} C_v &= 0.113Nk\left(\frac{kT}{2SJ}\right)^{3/2} \\ &= aT^{3/2}. \end{aligned} \tag{11.5}$$

8.11.2. Relation to Lattice Bose Gas

The Hamiltonian of Eq. (11.1) represents isotropic spin interactions. If there is anisotropy, the first term should be replaced by

$$H = -\sum_{(ij)} [S_{ix}S_{jx} + S_{iy}S_{jy} + (\Delta)S_{iz}S_{jz}], \tag{11.6}$$

where $\Delta = 1$ corresponds to isotropic ferromagnetism and $\Delta = -1$ to antiferromagnetism. Moreover, one can show that the anisotropic Heisenberg model has a relation to a lattice model of a Bose gas. In particular, the ground-state energy and the thermodynamic properties of a system with Eq. (11.6) yield the corresponding quantities for a Bose gas based on a lattice model.

In the lattice model, the quantum gas is assumed to have a kinetic energy expressed by a double difference instead of the standard operator $(-\hbar^2/2m)\nabla^2$. The energy of interactions between the nearest neighbors is equal to $-\Delta$. If $\Delta < 0$, the interaction is repulsive, and if $\Delta > 0$, it is attractive. Moreover, a hard core is assumed, to prevent two atoms from occupying the same lattice point.

The correspondence relations between the lattice Bose gas and a Heisenberg ferromagnet are tabulated in Table 8.2.

TABLE 8.2

Correspondence between a Lattice Quantum Gas and a Heisenberg Ferromagnet[a]

Lattice Bose gas	Heisenberg ferromagnet
Volume/cell size, v/v_0	Number of spins
Number density, n	$\langle S_z\rangle + \frac{1}{2}$
Chemical potential, μ	$H + [z/2(J-J')]$
pv_0	$-f + \frac{1}{2}H - \frac{1}{8}zJ'$

[a] In this table, v_0 represents the cell size, f is the free energy per lattice site, and z is the number of nearest-neighbor lattice points.

The correspondence relations were studied by Matsubara and Matsuda in terms of a molecular field theory (*35*). For a one-dimensional case, more

detailed analyses have been given. For instance, the lowest eigenvalue of the Hamiltonian for a given magnetization has been studied. (*36*).

As in the case of the Ising model, efforts have been made to find the time correlation functions. The behavior of these correlation functions is very crucial for understanding the dynamic properties of a Heisenberg magnet near the critical temperature. The problem is of course very difficult, but long-time correlations of spins are determined by small-frequency, i.e., near-static coupling. This limiting property and the behavior of the empirical laws near T_c give a clue for guessing the forms of correlation functions.

GENERAL REFERENCES

H. S. Green and C. A. Hurst, "Order-Disorder Phenomena." Wiley (Interscience), New York, 1964.

P. W. Kasteleyn, *in* "Fundamental Problems in Statistical Mechanics," Vol. II, p. 30 (E. G. D. Cohen ed.). Wiley, New York, 1968.

C. Domb, *in* " Magnetism," Vol. 2A (G. T. Rado and H. Suhl, eds.). Academic Press, New York, 1965.

R. Brout, *in* "Magnetism," Vol. 2A (G. T. Rado and H. Suhl, eds.). Academic Press, New York, 1965.

REFERENCES

1. E. Ising, *Z. Physik.* **31**, 253 (1925).

2. R. Peierls, *Proc. Camb. Philos. Soc.* **32**, 447 (1936).

3. H. A. Kramers and G. H. Wannier, *Phys. Rev.* **60**, 252 (1941).

4. W. Heisenberg, *Z. Physik.* **49**, 619 (1928).

5. E. W. Montroll, *Am. Math.Mo nthly* **61**, 46 (1954), *Proc. Third Berkeley Symp. Math. Stat.* p. 209 (1955).

6. G. F. Newell and E. W. Montroll, *Rev. Mod. Phys.* **25**, 353 (1953).

7. B. Kaufman, *Phys. Rev.* **76**, 1232 (1940).

8. T. D. Schultz, D. Mattis, and E. Lieb *Rev. Mod. Phys.* **36** 856 (1964).

9. M. Kac, P. L. Hemmer, and G. Uhlenbeck, *J. Math. Phys.* **4**, 216 (1963).

10. B. Zimm, P. Doty, and K. Iso, *Proc. Nat. Acad. Sci. U.S.A.* **45**, 160A (1959).

11. M. Kac and J. C. Ward, *Phys. Rev.* **88**, 1332 (1952).

12. L. Onsager, *Phys. Rev.* **65**, 117 (1944).

13. C. N. Yang, *Phys. Rev.* **85**, 808 (1952).

14. E. W. Montroll, R. B. Potts, and J. C. Ward, *J. Math. Phys.* **4**, 308 (1963).

15. L. P. Kadanoff, *Nuovo Cimento* **44**, 276 (1966).

16. T. D. Lee and C. N. Yang, *Phys. Rev.* **87**, 410 (1952).

17. R. Abe, *Progr. Theoret. Phys.* **38**, 72 (1967).

18. M. Suzuki, *Progr. Theoret. Phys.* **38**, 1224 (1967).

18a. A. Z. Patashinsky and V. L. Pokrovsky, *Soviet Physics JETP* **50**, 439 (1966) [English Transl. **23**, 292 (1966)].

18b. B. Widom, *J. Chem. Phys.* **43**, 3898 (1965).

19. L. Van Hove, *Phys. Rev.* **89**, 1189 (1953).

20. E. W. Montroll and R. B. Potts, *Phys. Rev.* **100**, 525 (1955).

21. K. Yamahuzi and T. Tanaka, *Progr. Theoret. Phys.* **20**, 327 (1958).

22. A. A. Maradudin, E. W. Montroll, and G. H. Weiss, "Theory of Lattice Dynamics in the Harmonic Approximation." Academic Press, New York, 1963.

23. A. E. Dixon, A. D. B. Woods, and B. N. Brockhouse, *Proc. Phys. Soc.* (*London*) **81**, 973 (1963).

24. G. Dolling *et al.*, *Phys. Rev.* **168**, 970 (1968).

25. J. L. Verble, J. L. Warren, and J. L. Yarnell, *Phys. Rev.* **168**, 980 (1968).

26. B. N Brockhouse *et al.*, *Phys. Rev.* **128**, 1099 (1962).

27. T. H. Berlin and M. Kac, *Phys. Rev.* **86**, 821 (1952).

28. A. Isihara, *Phys. Rev.* **108**, 619 (1957).

29. F. J. Dyson, *Phys. Rev.* **102**, 1217, 1230 (1956).

30. W. Opechowski, *Physica* **6**, 1112 (1938).

31. H. A. Brown and J. M. Luttinger, *Phys. Rev.* **100**, 685 (1955).

32. G. Horwitz and H. B. Callen, *Phys. Rev.* **124**, 1757 (1961).

33. N. D. Mermin and H. Wagner, *Phys. Rev. Lett.* **17**, 1133 (1966).

34. S. S. Shinozaki, *Phys. Rev.* **122**, 388 (1961).

35. T. Matsubara and H. Matsuda, *Progr. Theoret. Phys.* **16**, 569 (1956).

36. C. N. Yang and C. P. Yang, *Phys. Rev.* **150**, 327 (1966).

9

Phenomena near the Critical Temperature

The liquefaction of a gas cannot be achieved above the critical temperature. At this temperature, the local order of a liquid vanishes, and the difference in the properties of a liquid and its vapor disappears. Similar critical phenomena exist in many other systems. For instance, a magnet loses its ferromagnetic properties at the Curie temperature. Liquid helium loses its superfluidity at the λ-point.

When certain characteristic properties of a system disappear completely at a temperature, the term "second-order phase transition" is used. On the other hand, if a transition takes place portion by portion of a sample, it is "first-order."

It has been found that there exist similarities in different second-order phase transitions. Statistical properties of various substances follow strikingly simple laws near the critical temperatures. The statistical mechanics of the critical temperature phenomena have yet to be developed. It is precisely for this reason that this chapter is important.

9.1. Critical Temperature of a Fluid

The existence of a critical temperature was first found by Andrews in 1869 in work on CO_2. A few years later, van der Waals proposed his famous equation of state

$$p = \frac{kT}{v - b} - \frac{a}{v^2}. \tag{1.1}$$

The critical temperature T_c is found from the conditions

$$\frac{dp}{dv} = 0; \qquad \frac{d^2p}{dv^2} = 0; \qquad T = T_c. \tag{1.2}$$

Using these two conditions, we find from Eq. (1.1)

$$v_c = 3b; \qquad p_c = \frac{a}{27b^2}; \qquad kT_c = \frac{8a}{27b}. \tag{1.3}$$

The van der Waals equation of state played a historically important role in the theory of condensation. It was considered remarkable that a single equation could represent both the liquid and the gaseous phase. However, the isotherm according to Eq. (1.1) has an unstable part. Moreover, when the equation is actually calculated for CO_2 with reasonable values of atomic constants a and b, the pressure becomes negative in some region. Thus, only with the aid of *Maxwell's law* of equal areas does the equation of state become realistic to represent the condensation phenomenon. According to this law, the unstable part of the van der Waals isotherm is cut by a horizontal straight line to yield two areas of equal size between the line and the isotherm.

One important feature of the van der Waals equation is that the critical values can reduce the equation to

$$\begin{gathered} p^* = \frac{8T^*}{3v^* - 1} - \frac{3}{v^{*2}}, \\ p^* = \frac{p}{p_c}, \qquad v^* = \frac{v}{v_c}, \qquad T^* = \frac{T}{T_c}. \end{gathered} \tag{1.4}$$

The right-hand side is a universal function, and the van der Waals equation predicts the *law of corresponding states*. Kammerlingh Onnes was guided by this law in liquefying helium.

The law of corresponding states can be shown to hold for systems of the same potential function

$$\phi(r) = \phi_0 g\left(\frac{r}{r_0}\right), \tag{1.5}$$

where ϕ_0 and r_0 are constants characteristic of the given system. The Lennard-Jones potential has the form of Eq. (1.5). The configurational partition function can then be expressed in a reduced form:

$$Q(n^*, T^*) = Q\left(nr_0{}^3; \frac{kT}{\phi_0}\right), \tag{1.6}$$

The right-hand side is a universal function, n being the number density.

The law of corresponding states holds remarkably well for gas–liquid equilibria. Indeed, Guggenheim found an empirical formula (*1*)

$$\rho^* = 1 + \frac{3}{4}(1 - T^*) \pm \frac{7}{4}(1 - T^*)^{1/3} \tag{1.7}$$

for the density of the liquid phase (plus sign) and the gas phase (minus sign), ρ^* being the reduced density ρ/ρ_c. Up to the second term on the right-hand side of this expression, we have a rectilinear relation for the average density. The last term shows the $\frac{1}{3}$-*power law* for the density difference. On the other hand, the van der Waals equation predicts

$$v_g^* - v_l^* = 4(1 - T^*)^{1/2}, \tag{1.8}$$

that is, a $\frac{1}{2}$-power law.

Table 9.1 gives the actual values for T_c and p_c of some gases that obey Eq. (1.7).

TABLE 9.1

Critical Constants

	Ne	Ar	Kr	Xe	N_2	O_2	CO	CH_4
T_c (°K)	44.8	150.7	209.4	289.8	126.0	154.3	133.0	190.3
p_c (atm)	26.9	48.0	54.1	57.6	33.5	49.7	34.5	45.7

Figure 9.1 shows the density change $(\rho_l - \rho_g)/2\rho_c$ plotted against $[1 - (T/T_c)]^{1/3}$.

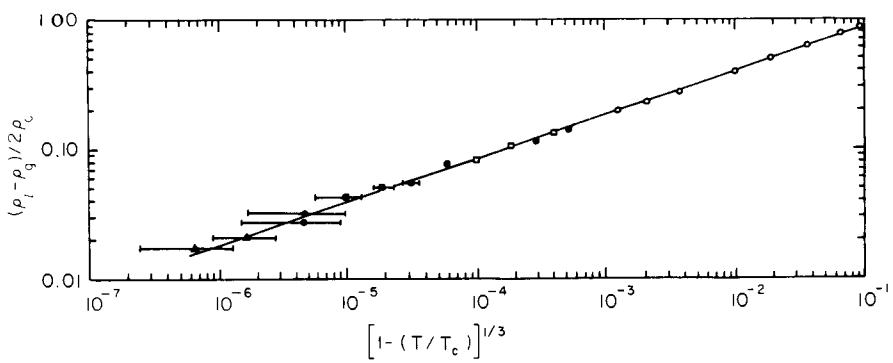


Fig. 9.1. The $\frac{1}{3}$-power law for the coexistence curve of CO_2. [After Heller (*2*).]

The power law holds remarkably well. On the other hand, for

$$\rho_l - \rho_g \propto (T_c - T)^{\beta} \tag{1.9}$$

with an adjustable parameter β, the following value has been reported for xenon:

$$\beta = 0.345 \pm 0.015.$$

Analogous to Eq. (1.9), it is possible to introduce a pressure–density relation:

$$|p - p_c| \propto |\rho - \rho_c|^{\delta}. \tag{1.10}$$

The exponent δ is approximately 4.3.

The measurement of specific heat near the critical temperature is not easy. According to recent investigations (*4*), the specific heats at constant volume of Ar and O_2 seem to diverge at T_c following the logarithmic law

$$c_V = -B^{\pm} \log\left|1 - \frac{T}{T_c}\right| + c^{\pm}, \tag{1.11}$$

where B^+ and B^- are of the same order of magnitude, and $c^- - c^+ \sim 10k$, $\pm$ corresponding to above and below T_c. The same logarithmic law has been found in the λ-transition of liquid helium and at some ferromagnetic

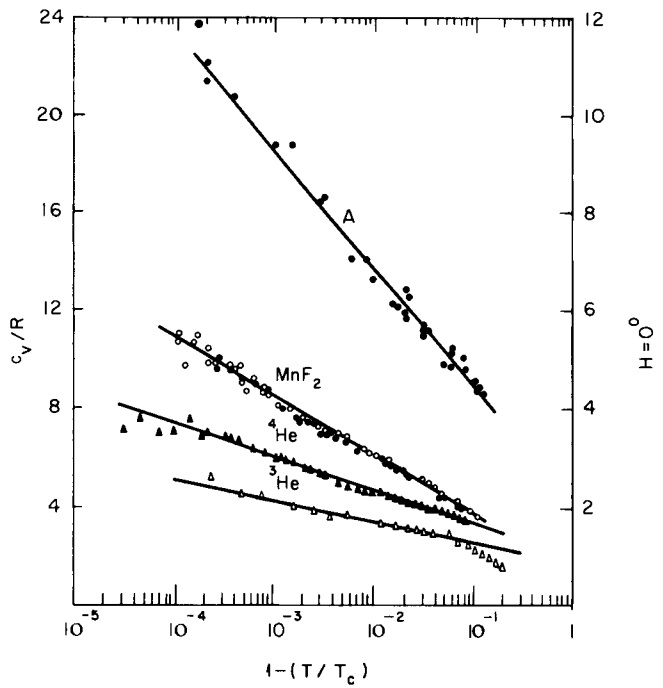


Fig. 9.2. The specific heats at constant volume for Ar, ^{4}He, ^{3}He (left ordinate), and antiferromagnet MnF_2 (right ordinate). [After Heller (*2*).]

transitions. The logarithmic law is in conformity with Onsager's rigorous solution of a two-dimensional Ising lattice. Figure 9.2 illustrates the specific heat data (*2*) for Ar, ^{4}He, ^{3}He, and MnF_2. It is remarkable that these different substances have the same logarithmic dependence.

9.2. Relationships among the Critical Exponents

The isothermal compressibility κ_T is defined by

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T. \tag{2.1}$$

At the critical temperature on the isotherm, κ_T diverges. If one uses a power law, the behavior can be expressed by

$$\kappa_T \propto |T - T_c|^{-\gamma}, \tag{2.2}$$

with $\gamma \approx 1.1$.

The compressibility power law (2.2) can be understood from a consideration based on a Taylor expansion (*5*). Let us expand the difference of the pressure from its critical value as follows:

$$\Delta p = a\,\Delta n + b\,\Delta T + c(\Delta n)^2 + d(\Delta n\,\Delta T) + e(\Delta T)^2 + f(\Delta n)^3 + \cdots, \tag{2.3}$$

where n is the density and Δ indicates the deviation from the critical point. Since $(\partial p/\partial V)_T = (\partial^2 p/\partial V^2)_T = 0$ at T_c, we take $a = c = 0$. The term $(\Delta T)^2$ is smaller than the term ΔT and may be neglected. The term $(\Delta n)^3$ must be kept because Δn and ΔT are independent of each other. Thus,

$$\Delta p = b\,\Delta T + d(\Delta n\,\Delta T) + f(\Delta n)^3 + \cdots \tag{2.4}$$

determines the pressure near the critical point.

In particular, on the critical isotherm, Eq. (2.4) leads to

$$\Delta p \propto (\Delta n)^3, \tag{2.5}$$

and

$$\kappa = \frac{1}{n}\left(\frac{\Delta n}{\Delta p}\right)_T \propto (\Delta T)^{-1}. \tag{2.6}$$

Equations (2.5) and (2.6) are in agreement with experimental laws given before. By a similar argument, one can find the relation

$$\delta = 1 + \frac{\gamma}{\beta}, \tag{2.7}$$

This relation is exactly satisfied by the van der Waals gas, for which $\beta = \frac{1}{2}$, $\gamma = 1$, and $\delta = 3$. However, experimental values deviate slightly from this prediction:

$$\delta \approx 4.3; \qquad \gamma = 1.1; \qquad \beta = 0.33\ldots. \tag{2.8}$$

Sometimes, a power law for the heat capacity is used:

$$c_V \propto (T_c - T)^{-\alpha}.$$

The value of α may differ above and below T_c. For $T < T_c$, this power α may be related to the powers β and γ as follows (5–8):

$$\alpha + 2\beta + \gamma = 2.$$

Problem

Using the thermodynamic inequality $\Delta F + p\,\Delta V > 0$, show that $(\partial^3 p/\partial V^3)_T < 0$, and then examine the sign of the coefficient d in Eq. (2.4). Here, F is the Helmholtz free energy.

9.3. Magnetic Phase Transitions

9.3.1. Ferromagnetism

The discovery of a critical temperature in a solid was made by Hopkinson. Later, in 1895, Pierre Curie confirmed the existence of the critical temperature and established the susceptibility–temperature relation by thorough experimentation. Since then, the name "Curie point" has been used for critical temperatures of magnetic and dielectric substances.

In the absence of a magnetic field, the spontaneous magnetization $M_0(T)$ near a ferromagnetic Curie point follows the empirical formula

$$M_0(T) \propto (T_c - T)^{\beta}, \qquad \beta = 0.33 \pm 0.03; \qquad T \to T_c -; \quad H = 0. \tag{3.1}$$

On the other hand, theoretical values are

$$\begin{aligned} &\beta = \tfrac{1}{2} && \text{(Weiss theory)},\\ &\beta = \tfrac{1}{8} && \text{(Ising square lattice theory)},\\ &\beta = 0.312 && \text{(Ising three-dimensional theory)}. \end{aligned} \tag{3.2}$$

For nickel and gadolinium, the critical isotherm, that is, the magnetization M plotted against H, follows the law

$$|M| \propto |H|^{1/\delta}, \tag{3.3}$$

where

$$\delta = 4.1 \pm 0.1, \qquad T = T_c; \quad H \to 0.$$

The relation (3.3) is analogous to the relation (1.10).

A magnetic analog to the isothermal compressibility of fluids is the isothermal susceptibility χ, which varies as

$$\chi \propto (T - T_c)^{-\gamma}, \qquad \gamma = 1.33 \pm 0.03; \qquad T \to T_c+; \quad H = 0. \tag{3.4}$$

The specific heat at constant magnetic field can be represented by

$$c_H \propto (T - T_c)^{-\alpha}, \qquad 0 \leqslant \alpha < 0.16; \qquad T \to T_c+; \quad H = 0. \tag{3.5}$$

or by a logarithmic function such as Eq. (1.11).

9.3.2. Antiferromagnetism

Antiferromagnetism appears when antispin configurations are favored energetically in comparison with parallel configurations (which are favorable in ferromagnetic states). The transition to an antiferromagnetic state from a paramagnetic state is characterized by a kink in the graph of susceptibility versus temperature. The transition point is called the *Néel temperature*, which is the critical point for this case.

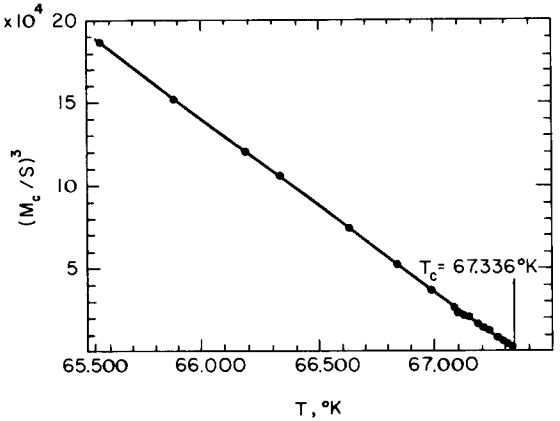


Fig. 9.3. The power law for antiferromagnet MnF_2. The cube of the zero-field nuclear magnetic resonance frequency is plotted against temperature. [After Heller and Benedek (*9*).]

A very accurate measurement of the magnetization of an antiferromagnet has been performed on MnF_2 (Fig. 9.3). This substance has the Néel temperature $T_c = 67.336°K$. In the case of antiferromagnetism, a sublattice magnetization is the quantity that corresponds to the magnetization of a ferromagnet. At zero magnetic field, the sublattice magnetization $M_0'(t)$ obeys the relation

$$M_0'(T) \propto (T_c - T)^{\beta}; \qquad \beta = 0.333 \pm 0.003. \tag{3.6}$$

This relation holds over a temperature range of $7 \times 10^{-5} < \Delta T/T_c < 0.1$. The closest point measured is about 3 mdeg from T_c (see Fig. 9.3).

The heat capacity of MnF_2 has been measured with less accuracy. The following relation seems to hold:

$$C \propto (T - T_c)^{-\alpha}, \qquad 0 \leqslant \alpha \leqslant 0.16; \qquad T \to T_c-. \tag{3.7}$$

According to careful measurements by Skalyo and Friedberg (*10*) on $CoCl \cdot 6H_2O$,

$$C_p(T)^{\pm} \sim R\left(\log\left|1 - \frac{T}{T_c}\right| \pm \Delta_{\pm}\right), \qquad H = 0, \tag{3.8}$$

where plus and minus correspond, respectively, to above and below the critical temperature and Δ is a constant. However, in the immediate vicinity of T_c, the same logarithmic dependence does not hold, presumably due to crystal strains and the coexistence of paramagnetic and antiferromagnetic phases.

9.4. Binary Mixtures

9.4.1. Gas–Gas Mixtures

In some binary gas mixtures, phase separations occur at temperatures well above the critical temperatures of their pure components. Phase separations can be expected when the molecules are complicated, but have also been observed in He–Xe systems (*11*). This is a remarkable phenomenon in view of the simple chemical structure of the constituent molecules. The critical temperature is a function of concentration and pressure. The critical locus, that is, the P_c versus T_c curve, is a straight line.

9.4.2. Liquid Solutions

Phase separations in liquid mixtures may take place above or below the (lower or upper) critical temperature. Some liquid mixtures have both upper and lower critical temperatures: Between these two temperatures, phase

separations take place. In contrast to the case of gas–liquid condensation, the role of pressure in the phase separations in liquid solutions is secondary. This makes experimental studies on critical phenomena in liquid solutions somewhat easier than in gases.

For this reason, there are many experiments on phase separations in liquid solutions. For mixtures of CCl_4 with n-C_7F_{16}, the phase separation obeys the $\frac{1}{3}$-power law

$$(x_1 - x_2) \propto (T_c - T)^{0.33}, \tag{4.1}$$

where x represents the mole fraction and the subscripts 1 and 2 correspond to the two phases.

9.4.3. Solid Solutions

Many binary alloys undergo an order–disorder phase change. One well-known example is that in β-brass, i.e., CuZn. This system shows a λ-type anomaly in the heat capacity versus temperature curve due to the relocation of constituent atoms from ordered to disordered positions. Similar anomalies have been found in Cu_3Au, FeNi, FeCo, and Ni_3Mn. X-ray and neutron diffraction experiments show the appearance of a superlattice structure below T_c.

Table 9.2 shows some typical power laws. The data in the upper half

TABLE 9.2

Comparison of Exponents for Various Power Laws

	$\rho_l - \rho_g \propto (T_c - T)^{\beta}$	$\kappa \propto (T_c - T)^{-\gamma}$	$p - p_c \propto (\rho - \rho_c)^{\delta}$
Ar	0.45 ± 0.07	—	4.2(+0.6, −0.3)
Xe	0.35 ± 0.015	1.3(+0.1, −0.2)	—
^{4}He	0.352 ± 0.015	—	—
^{3}He	0.45 ± 0.05	—	3.5 ± 0.1
H_2	0.375 ± 0.015	—	—
CO_2	0.34 ± 0.015	1.35 ± 0.15	5 ± 1
	$M_0 \propto (T_c - T)^{\beta}$	$\chi \propto (T_c - T)^{-\gamma}$	$H \propto M^{\delta}$
Fe	0.33 ± 0.015	1.333 ± 0.015	4.22 ± 0.1
Co	—	1.21 ± 0.04	—
Ni	0.365 ± 0.02	1.35 ± 0.02	—
CrO_2	—	1.62 ± 0.02	5.75 ± 0.05
MnF_2	0.335 ± 0.005		—

correspond to gas–liquid transitions and the data in the lower half to those in ferromagnets and antiferromagnets. Note that the powers are not necessarily the same in the same column.

9.5. Quantum Liquid Solution

The mixtures of liquid ^{3}He with liquid ^{4}He are very interesting because they represent a new set of Fermi liquids with different degeneracy temperatures. As the concentration of one of the components, say ^{3}He, is changed, one can study the transition from classical gaseous to quantum liquid behavior. Moreover, since the effective (attractive) interaction between ^{3}He quasi particles is considered to be weak, theoretical treatment becomes possible. Experimentally, the solution can be used for producing extremely low temperatures because the mixing energy is positive, i.e., heat is absorbed (^{3}He *refrigeration*).

Although ^{3}He atoms and ^{4}He atoms are similar to each other, liquid ^{3}He has a larger specific volume and a smaller internal energy than liquid ^{4}He, and their mixtures are not ideal solutions at all. In fact, below 0.87°K, the system separates into two phases (*12*), a heavier phase rich in ^{4}He, and a lighter one rich in ^{3}He. The phase separation has been observed visually, via sound velocity, density, specific heat, NMR, and slow neutron scattering.

In a phase-separated mixture of ^{3}He and ^{4}He, the chemical potential of each isotope must be the same in the upper and lower phases:

$$\begin{aligned} \mu_3(x_l, T) &= \mu_3(x_u, T), \\ \mu_4(x_l, T) &= \mu_4(x_u, T), \end{aligned} \tag{5.1}$$

where x is the molar concentration of ^{3}He, i.e., $x = N_3/(N_3 + N_4)$, where N_3 and N_4 represent the total numbers of ^{3}He and ^{4}He atoms, respectively. At very low temperatures, x and $1 - x_u$ become small, so that the ^{3}He atoms in the lower phase may be treated as impurities of very low concentrations.

Thus, the interactions between ^{3}He particles may be neglected in the first approximation. However, since they must push aside the ^{4}He atoms in their thermal movement, it is appropriate to introduce an effective mass m_3^*. The energy of a ^{3}He solute may be expressed by

$$\varepsilon_3 = -\varepsilon_3{}^0 + \frac{p^2}{2m_3^*}, \tag{5.2}$$

where $-\varepsilon_3{}^0$ is the binding energy. From the specific heat curve of the mixture, which looks like that of an ideal Fermi gas, it has been found that $m_3^* = 2.88m_3$, the mass of ^{3}He.

A quasi-particle theory of ^{3}He in superfluid ^{4}He was initiated by Landau and Pomeranchuk (*13*) in 1948. The theory has been improved by Emery (*14*) and by Bardeen *et al.* (*15*). The latter (BBP) theory takes into account the interactions between ^{3}He atoms by means of effective interactions. An important parameter in the BBP theory is the number $\alpha_0 = \alpha(x = 0, T = 0)$, which is defined in the molar volume expression of a solution of concentration x such that

$$v(x, T) = [1 + x\alpha(x, T)]v_4(T), \tag{5.3}$$

where v_4 is the molar volume of pure ^{4}He. Using

$$\left(\frac{\partial \mu_3}{\partial p}\right)_{x,T} = v + (1 - x)\left(\frac{\partial v}{\partial x}\right)_{p,T} \tag{5.4}$$

and $\mu_3 = -N_0 \varepsilon_0$ for 0°K and $x = 0$, we find

$$1 + \alpha_0 = -\frac{N_0}{v_4}\left(\frac{d\varepsilon_0}{dp}\right). \tag{5.5}$$

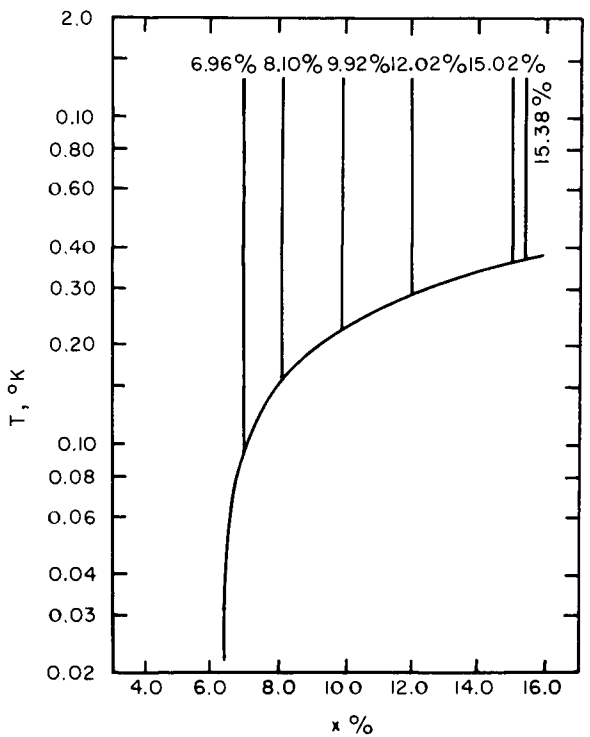


Fig. 9.4. The solubility curve for ^{3}He in ^{4}He. [After Edwards, Ifft, and Sarwinski (*17*).]

The right-hand side of Eq. (5.3) involves x in a complicated way. However, one can determine x and α by successive approximations starting with α_0. Edwards *et al.* (*17*) determined in this way the solubility curve, that is, the concentration of ^{3}He in the lower phase in the two-phase region [see also Edwards and Daunt (*18*)]. Figure 9.4 illustrates their data. Below 0.15°K, the data fit the empirical equation

$$x_l = x_0(1 + 10.8T^2), \tag{5.6}$$

where x_0, the solubility at 0°K, is $(6.40 \pm 0.07)\%$.

The upper phase is almost pure ^{3}He. The concentration X_u follows the empirical relation

$$1 - x_u = 1.13T^{3/2}e^{-0.71/T}, \qquad T < 0.4°\text{K}. \tag{5.7}$$

9.6. Order–Disorder Theory

Phase transitions from an ordered state to disordered state in alloys and magnets have been observed for many years and have been treated by various approximate methods.

The first approach to such *order–disorder phenomena* was made by Bragg and Williams in 1934. Since then, many refined theories have been developed by taking into account the more detailed local configurations of atoms. We shall, however, limit our consideration to the classical Bragg–Williams approximation because of its simplicity.

The phase transitions that are the results of interactions of particles or spins are often called *cooperative phenomena.*

9.6.1. Lattice Gas

A *lattice gas* is a theoretical model for a real gas. In this model, molecules can be located only on lattice sites. Let us consider N molecules in a lattice of $N + N'$ sites. Let us assume that only the molecules in the nearest-neighbor lattice sites interact with each other, and let $-\phi$ be the interaction potential.

If the molecules are distributed randomly, the number of pairs of nearest-neighbor molecules is $zN^2/2(N + N')$, where z is the lattice coordination number, that is, the number of nearest-neighbor lattice sites. The configurational energy of the lattice gas may be considered to be proportional to this number. Such an approximation based on a random mixing is due to Bragg and Williams. Later theories improve upon this point by introducing a parameter representing the *short-range order* of atoms. In the *Bragg–Williams approximation*, only the *long-range order* of atoms is taken into consideration.

The partition function of the lattice gas is given by

$$Z = \frac{(N+N')!}{N!\,N'!} \exp\left(\frac{z}{2}\frac{N^2}{N+N'}\frac{\phi}{kT}\right); \tag{6.1}$$

and $\beta = 1/kT$ and the combinatorial factor corresponds to a random mixing. From Eq. (6.1), we obtain the free energy F and the entropy S as follows:

$$\begin{aligned} F &= -\phi\frac{z}{2}\frac{N^2}{N+N'} + kT\left(N\log\frac{N}{N+N'} + N'\log\frac{N'}{N+N'}\right), \\ S &= -k\left(N\log\frac{N}{N+N'} + N'\log\frac{N'}{N+N'}\right). \end{aligned} \tag{6.2}$$

The first term in the free-energy expression is the internal energy. The second term comes from the combinatorial factor in the partition function and is called the mixing free energy. This term is obtained by making use of Stirling's formula. Here, S is the entropy of mixing.

To obtain the pressure, we introduce volume per lattice site v_0 such that

$$V = (N+N')v_0\,. \tag{6.3}$$

The pressure p due to a volume change

$$dV = v_0\,dN'$$

is given by

$$-p = \frac{1}{v_0}\left(\frac{z\phi}{2}\frac{V_0^2}{V^2} - kT\log\frac{V}{V-V_0}\right), \tag{6.4}$$

where $V_0 = Nv_0$ is the volume at $T = 0°\mathrm{K}$ (then there is no hole).

The critical constants are determined as follows:

$$kT_c = \tfrac{1}{4}z\phi, \qquad V_c = 2V_0\,. \tag{6.5}$$

With this simple treatment, the Kammerlingh–Onnes constant becomes

$$K = \frac{NkT_c}{p_c V_c} = 2.6. \tag{6.6}$$

This value is close to the experimental value of 3.3 for argon.

9.6.2. Ferromagnetism

In the preceding treatment, let N' be the number of up spins and N be the number of down spins. Let ε be the energy per opposite-spin pair in the nearest-neighbor sites. There are altogether $zNN'/(N+N')$ such pairs. The

free energy is

$$F = \varepsilon z\frac{NN'}{N+N'} + H(N-N') + kT\left(N\log\frac{N}{N+N'} + N'\log\frac{N'}{N+N'}\right), \tag{6.7}$$

where we have added the energy due to a magnetic field, expressed in a suitable unit.

The magnetic moment M and the intensity of magnetization I may be defined in our units by

$$M = N' - N = -\frac{\partial F}{\partial H}, \qquad I = \frac{N'-N}{N'+N}\,. \tag{6.8}$$

The number N can be found as a function of T by minimizing the free energy:

$$\left(\frac{\partial F}{\partial N}\right)_{N+N'=\text{const}} = 0. \tag{6.9}$$

Thus, we obtain

$$\frac{1}{2}\log\frac{1+I}{1-I} = \frac{1}{kT}H + \frac{\varepsilon z I}{2}. \tag{6.10}$$

By plotting both sides against I, it is found that there is no I that satisfies Eq. (6.10) when $T > T_c$, where

$$T_c = \frac{\varepsilon z}{2k}. \tag{6.11}$$

Therefore, T_c is the critical temperature.

The second term on the right-hand side of Eq. (6.10) may be interpreted as a molecular field. As in the case of the Weiss field, it is proportional to I. In other words,

$$H' = H + \tfrac{1}{2}\varepsilon z I \tag{6.12}$$

is the field acting on each molecular magnet. In this sense, the above treatment is equivalent to Weiss's molecular field treatment.

Unfortunately, the Bragg–Williams (BW) and Weiss theories do not reproduce sharp experimental phase transitions. As a matter of fact, at the phase-transition point, the BW theory gives a finite heat capacity anomaly:

$$\Delta C = \begin{cases} 0, & T > T_c\,, \\ \dfrac{3}{2}Nk\left(1 + \dfrac{8}{5}\dfrac{T-T_c}{T_c}\right), & T < T_c\,. \end{cases} \tag{6.13}$$

It also gives

$$I \propto \left(1 - \frac{T}{T_c}\right)^{1/2}; \qquad \chi \propto \left(1 - \frac{T}{T_c}\right)^{-1}. \tag{6.14}$$

The susceptibility exponent agrees with experiments, but in Eq. (6.14) the intensity of the magnetization has a power of $\frac{1}{2}$ instead of the experimental value of $\frac{1}{3}$.

9.6.3. Binary Alloys

Let us consider an alloy of $A + B$-type. Let x be the mole fraction of one of the components. The total nearest-neighbor interaction energy can be expressed by

$$\begin{aligned} U &= \tfrac{1}{2}Nz[x^2\varepsilon_{AA} + (1-x)^2\varepsilon_{BB} + 2x(1-x)\varepsilon_{AB}] \\ &= \tfrac{1}{2}Nz[x\varepsilon_{AA} + (1-x)\varepsilon_{BB} + 2x(1-x)\varepsilon], \end{aligned} \tag{6.15}$$

where the ε_{ij} represent the interaction energies per pair of ij-species and where

$$\varepsilon = \varepsilon_{AB} - \tfrac{1}{2}(\varepsilon_{AA} + \varepsilon_{BB}). \tag{6.16}$$

If $\varepsilon > 0$, A and B do not mix below a certain temperature. The phase separation takes place at

$$T_c = \frac{z\varepsilon}{4k}, \tag{6.17}$$

as in Eq. (6.11). When $\varepsilon < 0$, A and B attract each other in such a way that the alloy becomes a chemical compound. The entropy of mixing is found as in Eq. (6.2):

$$S = -Nk[x \log x + (1-x)\log(1-x)]. \tag{6.18}$$

Adding Eqs. (6.15) and Eq. (6.18), we obtain the free energy. Equation (6.17) follows from the free-energy expression.

As in the case of ferromagnetism, this theory does not yield a sharp phase transition for binary mixtures: The specific heat reaches a constant value and does not diverge.

9.7. Density Fluctuations near Critical Temperature

The fact that the compressibility of fluids diverges at the critical temperature suggests that the fluctuations of density become very large. The first theoretical treatment of the fluctuations was given by Ornstein and Zernike in 1917. Later, Landau and Lifshitz (*19*) presented a theory based on a Taylor expansion, which we shall discuss in this section.

Let us start with the density–density correlation function given by Eq. (7.20) of Chapter 6:

$$\langle n(\mathbf{r})n(\mathbf{r}')\rangle = n\,\delta(\mathbf{r}-\mathbf{r}') + \rho_2(|\mathbf{r}-\mathbf{r}'|), \tag{7.1}$$

where $n(\mathbf{r})$ is a local fluctuating number density and n is the average density. Using the difference $n(\mathbf{r}) - n$, one can rewrite Eq. (7.1) as follows:

$$\langle [n(\mathbf{r})-n][n(\mathbf{r}')-n]\rangle = n\,\delta(\mathbf{r}-\mathbf{r}') + \rho_2(|\mathbf{r}-\mathbf{r}'|) - n^2. \tag{7.2}$$

We assume that $n(\mathbf{r}) - n$ can be represented by a Fourier integral:

$$n(\mathbf{r}) - n = \frac{1}{(2\pi)^3}\int n_{\mathbf{k}} \exp i\mathbf{k}\cdot\mathbf{r}\, d\mathbf{k}. \tag{7.3}$$

Because $n(\mathbf{r}) - n$ is real, $n_{\mathbf{k}}^* = n_{-\mathbf{k}}$. Also, because of the normalization of $n(\mathbf{r})$, $n_0 = 0$. Equations (7.2) and (7.3) give

$$\langle |n_{\mathbf{k}}|^2\rangle = nV + V\int (\exp i\mathbf{k}\cdot\mathbf{r})[\rho_2(r) - n^2]\, d\mathbf{r}. \tag{7.4}$$

The statistical average on the left side may be found in the following way. We express the total free energy F in terms of the free-energy density $F_1(\mathbf{r})$:

$$F = \int F_1(\mathbf{r})\, d\mathbf{r}. \tag{7.5}$$

For a given temperature, $F_1(\mathbf{r})$ may deviate from its equilibrium value $\langle F_1\rangle$, due to the density fluctuations. Because $F_1(\mathbf{r})$ should be minimum for equilibrium, we expand the difference $F_1(\mathbf{r}) - \langle F_1\rangle$ as follows:

$$F_1(\mathbf{r}) - \langle F_1\rangle = \tfrac{1}{2}a[n(\mathbf{r}) - n]^2 + \tfrac{1}{2}b(\nabla n)^2 + \cdots. \tag{7.6}$$

where a and b are independent of the density but may depend on the temperature.

In Eq. (7.3), we have used a Fourier integral. For a finite volume, we should have used a Fourier series. The conversion can be made through

$$(2\pi)^{-3}\int d\mathbf{k} \to V^{-1}\sum_{\mathbf{k}}.$$

Using this relation and introducing Eq. (7.3) into Eq. (7.6) and neglecting higher-order terms, we find

$$F - \langle F_1\rangle V = \frac{1}{2V}\sum_{\mathbf{k}} (a + bk^2)|n_{\mathbf{k}}|^2. \tag{7.7}$$

Based on this expression, we assume that the probability of having $|n_{\mathbf{k}}|^2$ is given by

$$\frac{\exp[-(a+bk^2)|n_{\mathbf{k}}|^2/2VkT]}{\int \exp[-(a+bk^2)|n_{\mathbf{k}}|^2/2VkT]\, dn_{\mathbf{k}}}. \tag{7.8}$$

We use this probability function for evaluating $\langle |n_{\mathbf{k}}|^2 \rangle$. The result is

$$\langle |n_{\mathbf{k}}|^2 \rangle = \frac{VkT}{a + bk^2}. \tag{7.9}$$

Using this in Eq. (7.4) and taking the Fourier inverse, we obtain

$$\rho_2(r) - n^2 + n\,\delta(\mathbf{r}) = \frac{kT}{4\pi br} \exp\left[-\left(\frac{a}{b}\right)^{1/2} r\right] \qquad (b \neq 0). \tag{7.10}$$

Thus, the p.d.f. decreases exponentially if a is not zero.

The quantity a must vanish at the critical point because

$$a = \left(\frac{\partial^2 F_1}{\partial n^2}\right)_T = n^{-1}\left(\frac{\partial p}{\partial n}\right)_T.$$

Therefore, at the critical temperature, the p.d.f. becomes long-ranged:

$$\rho_2(r) - n^2 + n\delta(\mathbf{r}) = \frac{kT_c}{4\pi br}, \qquad T = T_c. \tag{7.11}$$

In the vicinity of T_c, if a varies as

$$a = a_0(T - T_c),$$

the compressibility is given by

$$\kappa = \frac{1}{n^2 kT} \int [\rho_2(r) - n\,\delta(\mathbf{r})]\, d\mathbf{r}$$

$$= \frac{1}{n^2 a_0} (T - T_c)^{-1}. \tag{7.12}$$

Namely, the compressibility of a fluid diverges in proportion to $(T - T_c)^{-1}$. This result is close to what has been found by experiments.

The above Landau theory is based on a Taylor series expansion. Therefore, its validity in the immediate vicinity of T_c must be questioned. According to Ginzburg (*20*) and others, it is valid for some small values of $|T - T_c|$ above a certain critical value determined by the range of interactions. Generally, the longer is the range of the forces, the closer to the critical temperature can the Landau theory be used. This might sound strange, but the classical van der Waals, Weiss, and Bragg–Williams theories are all based on long-range couplings of particles.

We summarize some typical experimental and theoretical power laws in Table 9.3. The same types of physical quantities are listed in the same row. The constants a and b depend on the type of the system, fluid or magnet. The powers for the three-dimensional Ising lattice are of course not exact. The two-dimensional Ising lattice is interesting because of the logarithmic specific heat, and the Heisenberg model of ferromagnetism gives the correct order of magnitude for γ.

TABLE 9.3

Comparison of Experimental and Theoretical Power Laws

Experiment		Theory			
Fluid system	Magnetic system	Classical	Ising 2D	Ising 3D	Heisenberg
$c_v \propto \log\lvert T_c - T\rvert$	—	$a\left(1 - b\dfrac{T_c - T}{T_c}\right)$	$\log(T_c - T)$	?	?
$\rho_l - \rho_g \propto (T_c - T)^\beta$	$M_0 \propto (T_c - T)^\beta$	$\frac{1}{2}$	$\frac{1}{8}$	0.313	—
$\beta \approx \frac{1}{3}$		—	—	—	—
$\kappa \propto \lvert T_c - T\rvert^{-\gamma}$	$\kappa_T \propto (T_c - T)^{-\gamma}$				
$\gamma \approx 1.2$?	$\gamma \approx 1.35$?	1	$1\frac{3}{4}$	1.250	$\sim 1\frac{1}{3}$
$\lvert p - p_c\rvert_{T_c} \propto \lvert \rho - \rho_c\rvert^\delta_{T_c}$	$\lvert H\rvert_{T_c} \propto \lvert M\rvert^\delta_{T_c}$	—	—	—	—
$\delta \approx 4.2$	$\delta \approx 4.22$	3	15	5.2	?

9.8. Spatial Correlation of a Bose Gas near the Condensation Temperature

The number density for an ideal Bose gas is given by Eq. (3.15) of Chapter 4, which in integral form is

$$n = \frac{1}{h^3}\int \frac{d\mathbf{p}}{\exp\{\beta[(p^2/2m) - \mu]\} - 1}. \tag{8.1}$$

The critical temperature corresponds to $\mu = 0$ and is determined from

$$n = \frac{(2\pi mkT_0)^{3/2}}{h^3} G_{3/2}(1), \tag{8.2}$$

where $G_{3/2}(1) = 2.612$. Just above T_0, μ is very small and one may expand $e^{-\beta\mu}$ in Eq. (8.1). We then find

$$n = n\left[1 + \frac{3}{2}\left(\frac{T - T_c}{T_c}\right)\right] - \frac{2\pi^{1/2}}{2.612} n(\beta|\mu|)^{1/2}, \tag{8.3}$$

which gives

$$\sqrt{(\beta|\mu|)} = 1.105\,\frac{T - T_c}{T_c} = \sqrt{\alpha}, \tag{8.4}$$

where we have defined α by this equation.

The pair distribution function for an ideal Bose gas has been obtained by London (*21*) and others (*22*, *23*) in the form (see Section 11.2, for the devriation)

$$\rho_2(r) = n^2 + \left[\frac{1}{h^3}\int \frac{\exp(i\mathbf{p}\cdot\mathbf{r}/h)}{\exp\{[\beta(p^2/2m) - \mu]\} - 1}\, d\mathbf{p}\right]^2. \tag{8.5}$$

Introducing Eq. (8.4) in Eq. (8.5), we arrive at

$$\frac{\rho_2(r)}{n^2} = (-3.54\alpha^{1/2} + 2.61 + 1.46\alpha)^{-2}\,\frac{\lambda^2}{r^2}\exp\left(-\frac{4\pi^{1/2}\alpha^{1/2}r}{\lambda}\right), \tag{8.6}$$

where λ is the thermal de Broglie wavelength.

Thus, the particle correlation function varies as $e^{-a(T-T_c)/T_c}/r^2$, similar to the cases of the Ornstein–Zernike and Landau theories, and becomes long-ranged as T approaches T_c.

The pair distribution function is related to the isothermal compressibility κ. It can easily be shown that, for an ideal Bose gas near the condensation point,

$$\kappa \propto (T - T_c)^{-1}. \tag{8.7}$$

The spatial correlation in liquid helium has been studied by neutrons and X rays. These measurements show that there is a change in the scattering pattern associated with the transition: The liquid below the λ-point has lower spatial order than just above it. According to a neutron experiment by Henshow (*24*), the principal maximum moves to larger angles and increases in height with increasing liquid density. The nearest distance between two helium atoms is found to be 2.27 ± 0.08 Å. The numbers of neighbors in the first and second coordinate shells change from about 6.5 to 8.5 atoms and from about 9 to 5.5 atoms, respectively, for density changes from 0.095 to 0.184 gm cm^{-3}, indicating a nearly close-packed structure.

9.9. Transport Coefficients near Critical Points

The study of transport phenomena near critical points is difficult because the system tends to move away from the critical region due to a nonequilibrium situation, such as a temperature gradient or a velocity gradient. In addition, the applicability of fundamental equations, such as the Navier–Stokes equation, and the linear phenomenological laws, such as Stokes's law, might not be valid in the critical region.

In spite of these difficulties, several different types of transport phenomena have been observed near critical points. Let us discuss some typical types.

(*i*) *Critical slowing-down*

Examples are the mutual diffusion constant of a critical solution and the inverse relaxation time of the ferroelectric salt $NaNO_2$.

(*ii*) *Critical divergence*

Examples are the thermal conductivity of CO_2 as a function of the density, the viscosity, and the thermal conductivity of liquids, and the width of the absorption curve in the electron and nuclear spin resonances of antiferromagnets.

(*iii*) *Critical dip*

An example is the observation that the diffusion constant of the binary alloy CuZn shows a kink at the critical point. The electrical conductivity of ferromagnets and the thermal conductivity of antiferromagnets dip at their critical temperatures.

Let us now give some theoretical considerations on the diffusion constant and the viscosity coefficient of a fluid. The diffusion constant of a binary

liquid mixture may be expressed in terms of the correlation of a current vector **J** (see Chapter 13):

$$D = \frac{1}{3\langle \Delta c^2 \rangle} \int_0^\infty dt \, \langle \mathbf{J}(t) \cdot \mathbf{J}(0) \rangle. \tag{9.1}$$

The square average concentration fluctuation is related to the derivative of the chemical potential by

$$N\langle \Delta c^2 \rangle = \frac{kT}{(\partial \mu / \partial c)_{p,T}}. \tag{9.2}$$

Therefore,

$$D = \frac{N}{3kT} \left(\frac{\partial \mu}{\partial c} \right)_{p,T} \int_0^\infty dt \, \langle \mathbf{J}(t) \cdot \mathbf{J}(0) \rangle. \tag{9.3}$$

At the critical temperature T_c,

$$\left(\frac{\partial \mu}{\partial c} \right)_T = 0, \tag{9.4}$$

and if the correlation remains finite, D may approach zero as T approaches T_c.

The viscosity coefficient is determined by the time correlation of the microscopic stress tensor in the following form:

$$\eta = \frac{1}{kT} \int_0^\infty dt \, d\mathbf{r} \, \langle J_{xy}(\mathbf{r}, t) J_{xy}(0, 0) \rangle. \tag{9.5}$$

Thus, the correlation function of the stress tensor must be divergent at the critical temperature in order to explain the experimental data (*25*, *26*).

REFERENCES

1. E. A. Guggenheim, *J. Chem. Phys.* **13**, 255 (1945).
2. P. Heller, *Rep. Progr. Phys.* **30**, 731 (1967).
3. M. S. Green and J. V. Sengers (eds.) "Critical Phenomena," p. 21. National Bureau of Standards, Washington, D.C., 1966.
4. M. I. Bagatskii, A. V. Veronel, and B. G. Gusak, *Soviet Physics JETP* **16**, 517 (1963).
5. L. Landau and E. M. Lifshitz, "Statistical Physics," p. 265. Addison-Wesley, Reading, Massachusetts, 1958.
6. M. E. Fisher, *J. Math. Phys.* **5**, 944 (1964).
7. R. B. Griffiths, *Phys. Rev.* **158**, 176 (1967).
8. B. Widom, *J. Chem. Phys.* **43**, 3898 (1965).
9. P. Heller and G. B. Benedek, *Phys. Rev. Lett.* **8**, 428 (1962).
10. J. Skalyo, Jr. and S. A. Friedberg, *Phys. Rev. Lett.* **13**, 133 (1964).
11. J. DeSwaan Arons and G. A. M. Diepen, *J. Chem. Phys.* **44**, 2322 (1966).

12. I. Pomeranchuk, *J. Exp. Theoret. Phys. (USSR)* **19**, 42 (1949).
13. D. Landau and I. Pomeranchuk, *Dokl. Akad. Nauk SSSR* **59**, 669 (1949).
14. V. J. Emery, *Phys. Rev.* **161**, 194 (1967).
15. J. Bardeen, G. Baym, and D. Pines, *Phys. Rev.* **156**, 207 (1967).
16. G. Baym and C. Ebner, *Phys. Rev.* **170**, 346 (1968).
17. D. O. Edwards, E. M. Ifft, and R. E. Sarwinski, *Phys. Rev.* **177**, 380 (1969).
18. D. O. Edwards and J. G. Daunt, *Phys. Rev.* **124**, 640 (1961).
19. L. Landau and E. M. Lifshitz, "Statistical Physics," p. 366. Addison-Wesley, Reading, Massachusetts, 1958.
20. V. L. Ginzburg, *Fiz. Tverd. Tela* **2**, 2031. [English Trans.: *Soviet Physics Solid State* **2**, 1824 (1960).]
21. F. London, *J. Chem. Phys.* **11**, 203 (1943).
22. A. Isihara, *J. Phys. Soc. Japan Suppl.* **26**, 256 (1958).
23. M. Luban and M. Revzen, *J. Math. Phys.* **9**, 347 (1968).
24. F. G. Henshow, *Phys. Rev.* **119**, 14 (1960).
25. K. Kawasaki, *Phys. Rev.* **150**, 291 (1966).
26. M. Fixman, *in* "Advances in Chemical Physics," Vol. VI (I. Prigogine, ed.). Wiley (Interscience), New York, 1964.

PART

III

QUANTUM INTERACTING SYSTEMS

10

Propagator Methods for the Partition Functions

In this chapter, we shall present a propagator formalism for the partition function based on the Bloch equation for the density matrix. The propagator is in mathematical language called the Green's function and is essentially the density matrix in physical terminology.

The propagator, its resolvent, and other formal aspects of the Bloch equation were discussed by Husimi (1) in his thesis in 1940. The name "propagator" and a diagram representation appeared in Feynman's famous positron theory of 1949. An application of the same method to statistical mechanical problems was first given by Goldberger and Adams (2) and by Siegert (3) in 1952. Since around 1955, quantum many-body theories based on graphs have made remarkable strides due to many efforts in many countries (4–10, 12, 13).

We shall discuss first the basic properties of the density matrix and then evaluate the equation of state or the energy of a quantum mechanical many-body system.

The graphs in this chapter are not labeled, while those in the next chapter on the distribution functions are labeled. These two chapters correspond to Chapters 5 and 6 on classical systems.

10.1. Density Matrix

In quantum mechanics, the observed value of a physical quantity of a system is given by the expectation value of the operator $\mathbf{G}$ for the physical quantity. If the system is described by a single wave function $\psi(\mathbf{r}, t)$, the expectation value is

$$\overline{\mathbf{G}} = (\psi, \mathbf{G}\psi). \tag{1.1}$$

The wave function $\psi(\mathbf{r}, t)$ may be expanded in terms of a complete orthonormal set of functions ϕ_n:

$$\psi(\mathbf{r}, t) = \sum_n c_n(t)\phi_n(\mathbf{r}), \tag{1.2}$$

where it is assumed that the $\phi_n(\mathbf{r})$ are independent of time. Then,

$$\overline{\mathbf{G}} = \sum_{n,m} c_m^* c_n (m|\mathbf{G}|n). \tag{1.3}$$

We introduce a projection operator P by

$$(n|P|m) = c_n c_m^*, \tag{1.4}$$

so that

$$\overline{\mathbf{G}} = \sum_{n,m} (n|P|m)(m|\mathbf{G}|n) = \operatorname{tr} P\mathbf{G}. \tag{1.5}$$

The name projection operator is due to the property

$$P^2 = P. \tag{1.6}$$

In addition, P has the property

$$(n|P|m) = (m|P|n)^*, \tag{1.7}$$

that is, P is a Hermitian operator.

If a system cannot be described by a single wave function, an ensemble of similar systems can be introduced to obtain an ensemble average as in the classical cases. We define the ensemble average of $\mathbf{G}$ by

$$\langle\overline{\mathbf{G}}\rangle = \sum_{n,m} \langle c_n c_m^*\rangle (m|\mathbf{G}|n). \tag{1.8}$$

Thus, it is convenient to introduce the *density matrix* ρ by the following elements:

$$(n|\rho|m) = \langle c_n c_m^*\rangle \tag{1.9}$$

and write the ensemble average as follows:

$$\langle\overline{\mathbf{G}}\rangle = \operatorname{tr} \rho\mathbf{G}. \tag{1.10}$$

Note that double averages enter here: One is purely quantum mechanical and the other is statistical. While both are essential to the average values, we shall simplify the notation by omitting the bar. In any case, a system is called *pure* if ρ reduces to P. Otherwise, it is *mixed.*

The density matrix has several important properties. First, it is Hermitian:

$$(n|\rho|m) = (m|\rho|n)^*. \tag{1.11}$$

In particular, the diagonal elements of a density matrix are real. If $\mathbf{G} = 1$ in Eq. (1.10), we find

$$\operatorname{tr} \rho = 1. \tag{1.12}$$

Thus, we conclude that the diagonal element $(n|\rho|n)$ of a density matrix gives the normalized probability that the state ϕ_n is realized in the ensemble.

Second, the averages defined by density matrices are invariant under unitary transformations. Suppose we introduce a new representation $\{\chi_n\}$ such that

$$\chi = \phi \cdot S, \tag{1.13}$$

where S is an unitary transformation:

$$S \cdot (S')^* = 1. \tag{1.14}$$

$(S')^*$ is the adjoint matrix of S. We observe that

$$\psi = \sum a_t \chi_t, \qquad a_t = \sum_n c_n (S^{-1})_{tn}, \tag{1.15}$$

and

$$(s|\rho'|t) = a_s a_t^* = \sum_{n,m} (S^{-1})_{sn} c_n c_m^* S_{mt}. \tag{1.16}$$

That is,

$$\rho' = S^{-1}\rho S. \tag{1.17}$$

If we evaluate the average of a physical quantity $\mathbf{G}'$ in the new representation $\{\chi_n\}$ in terms of the new density matrix ρ', we obtain

$$\begin{aligned} \operatorname{tr} \mathbf{G}'\rho' &\equiv \operatorname{tr} S^{-1}\mathbf{G}S \cdot S^{-1}\rho S \\ &= \operatorname{tr} \mathbf{G}\rho, \end{aligned} \tag{1.18}$$

that is, the same value as in the old representation. This invariant property makes the averaging operation in terms of the density matrix meaningful.

The time dependence of the wave function ψ is given by the Schrödinger equation

$$-\frac{\hbar}{i}\frac{\partial\psi}{\partial t} = H\psi. \tag{1.19}$$

The expansion (1.2) gives

$$-\frac{\hbar}{i}\sum_n \frac{dc_n}{dt}\phi_n = \sum_n c_n H\phi_n. \tag{1.20}$$

Using Eq. (1.20) in Eq. (1.9), one can show

$$\frac{d}{dt}(n|P|m) = \frac{i}{\hbar}(n|PH - HP|m). \tag{1.21}$$

In terms of the density matrix ρ and the commutator [,], Eq. (1.21) yields

$$\frac{\partial \rho}{\partial t} = \frac{i}{\hbar}[\rho, H]. \tag{1.22}$$

This is the quantum mechanical analog of the classical Liouville equation and is called the *Bloch equation*. The formal solution of Eq. (1.22) for a time-independent H is

$$\rho(t) = [\exp(-i\hbar^{-1}Ht)]\rho(0)\exp(i\hbar^{-1}Ht). \tag{1.23}$$

In particular, for systems in equilibrium, ρ shall be independent of time and

$$[\rho, H] = 0. \tag{1.24}$$

That is, the density matrix for this case commutes with the Hamiltonian. From the analogy to classical cases, we introduce an operator

$$\begin{aligned}\rho &= e^{-\beta H} \\ &= 1 - \beta H + \tfrac{1}{2}\beta^2 H^2 - \cdots\end{aligned} \tag{1.25}$$

as the density matrix for the canonical ensemble, β being $1/kT$.

In terms of the density matrix, a "phase-space distribution function" may be introduced. It is often called the *Wigner distribution function* and is defined by (*11*)

$$f(\mathbf{r}, \mathbf{v}; t) = \left(\frac{m}{h}\right)^{3N}\int \rho_N\left(\mathbf{r} - \frac{\mathbf{r}'}{2}, \mathbf{r} + \frac{\mathbf{r}'}{2}; t\right)\exp(im\hbar^{-1}\textstyle\sum \mathbf{r}_i' \cdot \mathbf{v}_i)\, d\mathbf{r}'^N \tag{1.26}$$

for a system of N particles. The right-hand side can be expanded in powers of h^2. In the limit of $\hbar \to 0$, the phase-space distribution function reduces to the classical velocity distribution function.

10.2. Density Matrix in the Canonical Ensemble

We consider a system of N fermions or bosons described by the Hamiltonian H. As we discussed in Chapter 4, quantum ideal gases show properties which are different from those of classical gases, especially at low temperatures and high densities. For the present, we are interested in quantum mechanical systems with interactions.

We shall assume that our system is in equilibrium and the Hamiltonian operator H has the eigenvalues E_n and the eigenfunctions ψ_n:

$$H\psi_n = E_n\psi_n. \tag{2.1}$$

The partition function of this system is

$$Z_N = \sum_n e^{-E_n/kT}. \tag{2.2}$$

The evaluation of the partition function requires two steps, first, to solve the Schrödinger equation, and then, to sum the series in Eq. (2.2). In general, the first step, which is not statistical mechanical but quantum mechanical, may not be possible: The Schrödinger equation is not always soluble. Therefore, we try to express the partition function directly in terms of the Hamiltonian. For this purpose, let us start with the operator introduced in Eq. (1.25):

$$e^{-\beta H} = 1 - \beta H + \frac{1}{2}\beta^2 H^2 - \cdots, \qquad \beta = \frac{1}{kT}; \tag{2.3}$$

and define

$$\begin{aligned}\rho(\mathbf{r}'^N, \mathbf{r}^N) &= \sum_n \psi_n^*(\mathbf{r}'^N)e^{-\beta E_n}\psi(\mathbf{r}^N) \\ &= \sum_n \psi_n^*(\mathbf{r}'^N)e^{-\beta H}\psi_n(\mathbf{r}^N),\end{aligned} \tag{2.4}$$

where the summation is over all the eigenstates and $\mathbf{r}^N$ represents the coordinates of the N particles. Here, $\rho(\mathbf{r}'^N, \mathbf{r}^N)$ is the $(\mathbf{r}'^N, \mathbf{r}^N)$ element of the density matrix ρ in the canonical ensemble. The trace of this matrix is simply the partition function:

$$\begin{aligned}\int \rho(\mathbf{r}'^N, \mathbf{r}^N)\,\delta(\mathbf{r}'^N, \mathbf{r}^N)\,d\mathbf{r}'^N\,d\mathbf{r}^N &= \int \rho(\mathbf{r}^N, \mathbf{r}^N)\,d\mathbf{r}^N \\ &= \sum_n e^{-\beta E_n},\end{aligned} \tag{2.5}$$

because of the normalization condition of ψ_n.

In Eq. (2.4), the density matrix is defined in terms of the energy eigenfunctions. However, it is possible to use any orthonormal and complete set of eigenfunctions to express ρ. To see this important point, let us expand the energy eigenfunctions in terms of the orthonormal set $\{\varphi_l\}$:

$$\psi_n = \sum_l c_{nl}\varphi_l, \qquad c_{nl} = \int \psi_n(\mathbf{r}^N)\varphi_l^*(\mathbf{r}^N)\,d\mathbf{r}^N. \tag{2.6}$$

Then,

$$\rho(\mathbf{r}'^N, \mathbf{r}^N) = \sum_n \sum_l \left(c_{nl}\varphi_l(\mathbf{r}'^N)\right)^* e^{-\beta H}\sum_s c_{ns}\varphi_s(\mathbf{r}^N). \tag{2.7}$$

Using the closure property of ψ_n,

$$\sum_n \psi_n^*(\mathbf{r}'^N)\psi_n(\mathbf{r}^N) = \delta(\mathbf{r}^N - \mathbf{r}'^N), \tag{2.8}$$

and the orthonormal property of φ_s, we find

$$\sum_n c_{nl}\, c_{ns} = \delta_{ls}. \tag{2.9}$$

Thus, Eq. (2.7) is reduced to

$$\rho(\mathbf{r}'^N, \mathbf{r}^N) = \sum_l \varphi_l^*(\mathbf{r}'^N)e^{-\beta H}\varphi_l(\mathbf{r}^N), \tag{2.10}$$

which is in the same form as Eq. (2.4).

In summary, the density matrix for thermal equilibrium may be defined in terms of any orthonormal, complete set of functions as in Eq. (2.10), and the partition function can be obtained by taking its trace.

In addition to this useful property of the trace, the diagonal elements of the density matrix for thermal equilibrium have an important meaning. If we write the diagonal element $\rho(\mathbf{r}^N, \mathbf{r}^N) = f(\mathbf{r}^N)$, then

$$\frac{f(\mathbf{r}^N)}{Z_N}\, d\mathbf{r}^N = \frac{N!}{Z_N}\sum_n \psi_n^*(\mathbf{r}^N)e^{-\beta H}\psi_n(\mathbf{r}^N)\, d\mathbf{r}^N \tag{2.11}$$

is the probability of finding the N particles in the configuration specified by $\mathbf{r}^N$ at temperature T. Using this probability, we can evaluate the averages of functions in the coordinate space.

Upon differentiating Eq. (2.4), we find

$$\frac{\partial \rho}{\partial \beta} = -H\rho, \tag{2.12}$$

which is the *Bloch equation* for the equilibrium density matrix. The density matrix is obtained as a solution of this Schrödinger-type differential equation. The formal solution can be expressed as follows:

$$\rho = e^{-\beta H}\rho_0,$$

where ρ_0 is the value of ρ at $\beta = 0$. From the normalization of ψ, we find

$$\rho_0 = \frac{1}{N!}\begin{vmatrix} \delta(\mathbf{r}_1' - \mathbf{r}_1), \ldots, \delta(\mathbf{r}_1' - \mathbf{r}_N) \\ \vdots \qquad\qquad \vdots \\ \delta(\mathbf{r}_N' - \mathbf{r}_1), \ldots, \delta(\mathbf{r}_N' - \mathbf{r}_N) \end{vmatrix}_{\pm}. \tag{2.13}$$

Here, the minus sign indicates determinant for fermions and the plus sign corresponds to permanent for bosons, a permanent being obtained by replacing all the minus signs of the corresponding determinant by plus signs. If Boltzmann statistics is used, ρ_0 is simply

$$\rho_0 = \delta(\mathbf{r}_1' - \mathbf{r}_1)\,\delta(\mathbf{r}_2' - \mathbf{r}_2)\cdots\delta(\mathbf{r}_N' - \mathbf{r}_N). \tag{2.14}$$

10.3. Simple Examples of the Density Matrix

We shall give two simple examples of the density matrix for equilibria.

10.3.1. Ideal Quantum Gases

The wave function may be expressed as

$$\psi_{\mathbf{k}} = \frac{1}{(N!\,V^N)^{1/2}}\sum_P (\pm)^P \exp(i\mathbf{k}^N \cdot P\mathbf{r}^N), \tag{3.1}$$

where $\mathbf{k}^N = (\mathbf{k}_1, \mathbf{k}_2, \ldots, \mathbf{k}_N)$ and $\mathbf{r}^N = (\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$. The upper and lower signs correspond, respectively, to bosons and fermions and P is a symbol to represent permutations. The density matrix is

$$\begin{aligned}\rho(\mathbf{r}^N, \mathbf{r}'^N) &= \frac{1}{N!^2 V^N}\sum_{\mathbf{k}}\sum_P\sum_{P'} \exp\left[-\frac{\beta\hbar^2}{2m}(\mathbf{k}^N \cdot \mathbf{k}^N)\right] \\ &\quad \times \exp[i\mathbf{k}^N \cdot (P\mathbf{r}^N - P'\mathbf{r}'^N)] \\ &= \frac{1}{N!}\sum_P (\pm)^P \prod_i \rho(\mathbf{r}_i, P\mathbf{r}_i').\end{aligned} \tag{3.2}$$

In the first expression, one $N!$ comes from $\psi_{\mathbf{k}}\psi_{\mathbf{k}}^*$. If the $\mathbf{k}$ sums are performed independently, another $N!$ is necessary because each state is invariant under interchange of N $\mathbf{k}$'s. In the last expression, the double permutations P and P' are replaced by a single one and, correspondingly, one $N!$ in the denominator drops out. The ρ's on the right-hand side represent the single-particle density matrices. More explicitly, Eq. (3.2) is

$$\rho(\mathbf{r}^N, \mathbf{r}'^N) = \frac{1}{N!\,\lambda^{3N}}\sum_P (\pm)^P \exp\left[\sum_i -\frac{\pi}{\lambda^2}(\mathbf{r}_i - P\mathbf{r}_i')^2\right], \tag{3.3}$$

where λ is the thermal de Broglie wavelength, given by

$$\lambda^2 = \frac{h^2}{2\pi m k T}. \tag{3.4}$$

Taking the diagonal elements, we obtain the distribution function:

$$f(\mathbf{r}^N) = \frac{1}{N!\lambda^{3N}} \sum_P (\pm)^P \exp\left[\sum_i -\frac{\pi}{\lambda^2}(\mathbf{r}_i - P\mathbf{r}_i)^2\right]. \tag{3.5}$$

Upon integrating $f(\mathbf{r}^N)$, the partition function can be derived.

10.3.2. One-Dimensional Harmonic Oscillator

The Hamiltonian and its eigenvalues are

$$H = -\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + \frac{\kappa}{2}x^2,$$

$$\varepsilon_n = (n + \tfrac{1}{2})\hbar\omega,$$

where

$$\omega = \left(\frac{\kappa}{m}\right)^{1/2}.$$

We also define $a^2 = (m\kappa)^{1/2}/\hbar$. Then, the eigenfunctions are

$$\psi_n(x) = \left(\frac{a}{\pi^{1/2}2^n n!}\right)^{1/2} \exp\left(\frac{-a^2x^2}{2}\right) H_n(ax). \tag{3.6}$$

Accordingly, the density matrix is given by

$$\rho(x, x') = \frac{a}{\pi^{1/2}} \sum_{n=0}^{\infty} \left\{\exp\left[-\beta\left(n + \frac{1}{2}\right)\hbar\omega\right]\right\} \frac{1}{(2^n)n!} \times H_n(ax)H_n(ax') \exp[-a^2(x^2 + x'^2)]. \tag{3.7}$$

In these expressions,

$$H_n(x) = (-)^n(\exp x^2)\frac{d^n(\exp -x^2)}{dx^n}$$

is the Hermite polynomial. If we use the Fourier transform relation

$$\exp(-x^2) = \frac{1}{\pi^{1/2}} \int_{-\infty}^{\infty} \exp(-u^2 + 2ixu)\, du, \tag{3.8}$$

we obtain the expression

$$H_n(x) = (-)^n(\exp x^2)\frac{(2i)^n}{\pi^{1/2}} \int_{-\infty}^{\infty} u^n \exp(-u^2 + 2ixu)\, du. \tag{3.9}$$

Introducing this into Eq. (3.7), we get the density matrix

$$\begin{aligned}
\rho(x, x') &= \frac{a\exp(-\beta\hbar\omega/2)}{\pi^{3/2}} \left[\exp\frac{a^2(x^2 + x'^2)}{2}\right] \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \sum_{n=0}^{\infty} \frac{1}{n!} 2^n \\
&\quad \times [-(\exp -\beta\hbar\omega)uv]^n \exp[-u^2 - v^2 + 2ia(xu + x'v)]\, du\, dv \\
&= \frac{a\exp(-\beta\hbar\omega/2)}{\pi^{3/2}} \left[\exp\frac{a^2(x^2 + x'^2)}{2}\right] \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp[-u^2 - v^2 \\
&\quad + 2ia(xu + x'v) - 2uv\exp -\beta\hbar\omega]\, du\, dv \\
&= \frac{a}{(2\pi\sinh\beta\hbar\omega)^{1/2}} \exp\left\{-\frac{1}{4}\left[\tanh\left(\frac{\beta\hbar\omega}{2}\right)\right]a^2(x + x')^2 \right. \\
&\quad \left. -\frac{1}{4}\left[\coth\left(\frac{\beta\hbar\omega}{2}\right)\right]a^2(x - x')^2\right\}.
\end{aligned} \tag{3.10}$$

At high temperatures, this becomes

$$\rho(x, x') \sim \frac{a}{(2\pi\beta\hbar\omega)^{1/2}} \exp\left[-a^2\frac{(x - x')^2}{2\beta\hbar\omega}\right], \tag{3.11}$$

which approaches $\delta(x - x')$ as $\beta \to 0$. On the other hand, at low temperatures, we have

$$\rho(x, x') \sim \frac{a}{[\pi\exp(\beta\hbar\omega)]^{1/2}} \exp\left[-\frac{a^2}{2}(x^2 + x'^2)\right]. \tag{3.12}$$

The probability that the one-dimensional harmonic oscillator particle is at x is given by

$$f(x) = \frac{a}{(2\pi\sinh\beta\hbar\omega)^{1/2}} \exp\left\{-\tanh\left(\frac{\beta\hbar\omega}{2}\right)a^2x^2\right\}. \tag{3.13}$$

The partition function is

$$Z = \int_{-\infty}^{\infty} f(x)\, dx = \frac{e^{\beta\hbar\omega/2}}{e^{\beta\hbar\omega} - 1}, \tag{3.14}$$

which coincides with Eq. (1.3) of Chapter 4.

At high temperature, the probability given by Eq. (3.13) becomes

$$f(x) = \frac{a}{(2\pi\beta\hbar\omega)^{1/2}} \exp\left(-\frac{\beta}{2}\kappa x^2\right), \tag{3.15}$$

which is simply the Boltzmann factor because $\kappa x^2/2$ is the potential energy.

Using Eq. (3.13), we can evaluate the average potential energy. For this purpose, we take the average of x^2:

$$\langle x^2\rangle = Z^{-1}\int_{-\infty}^{\infty} x^2 f(x)\, dx = \frac{1}{2a^2}\coth\frac{1}{2}\beta\hbar\omega. \tag{3.16}$$

Examination of this expression is left to the reader.

10.4 Propagator in the r–β Space

Let us consider a system with the Hamiltonian

$$H = H_0 + \Phi; \qquad \Phi = \sum_{ij} \phi_{ij}. \tag{4.1}$$

In view of Eq. (2.4), we introduce a new wave function in the $\mathbf{r}$–β space (i.e., space of coordinates and reciprocal thermal energy) by (*13*)

$$\begin{aligned} \Psi(\mathbf{r}^N; \beta) &= e^{-\beta H}\psi_n(\mathbf{r}^N) \\ &= e^{-\beta E_n}\psi_n(\mathbf{r}^N). \end{aligned} \tag{4.2}$$

Since H and $e^{-\beta H}$ commute, $\Psi_n(\mathbf{r}^N, \beta)$ satisfies the Bloch equation

$$\frac{\partial \Psi_n}{\partial \beta} = -H\Psi_n. \tag{4.3}$$

The solution of this equation may be obtained by solving the integral equation (*2*)

$$\Psi_n(\mathbf{r}^N, \beta) = \int K(\mathbf{r}^N\beta; \mathbf{r}'^N\beta')\Psi_n(\mathbf{r}'^N, \beta')\, d\mathbf{r}'^N, \tag{4.4}$$

as can easily be verified by differentiating both sides with respect to β. Note that the kernel should have the same symmetry property as $\Psi_n(\mathbf{r}^N, \beta)$ with respect to exchange of the coordinates $\mathbf{r}^N$.

The kernel represents a propagation of the set of coordinates from $(\mathbf{r}'^N, \beta')$ to $(\mathbf{r}^N, \beta)$ and is called the *propagator*. It is actually the Green's function satisfying

$$\frac{\partial K(2, 1)}{\partial \beta_2} + H(2)K(2, 1) = \delta(2 - 1). \tag{4.5}$$

Here, for simplicity, we used the notation 2 for $(\mathbf{r}^N, \beta = \beta_2)$ and 1 for $(\mathbf{r}'^N, \beta' = \beta_1)$. More explicitly, Eq. (4.4) can be rewritten as

$$\psi_n(\mathbf{r}^N) = \{\exp[(\beta - \beta')E_n]\} \int K(\mathbf{r}^N\beta; \mathbf{r}'^N\beta')\psi_n(\mathbf{r}'^N)\, d\mathbf{r}'^N. \tag{4.6}$$

From this expression, we see that $\psi_n(\mathbf{r}^N)$ is an eigenfunction and $\lambda_n = e^{(\beta-\beta')E_n}$ is an eigenvalue of the integral equation (4.6).

Equation (4.5) requires that the kernel be the Green's function belonging to the differential equation (4.3). The Green's function is subject to the subsidiary condition

$$\begin{aligned} K(2, 1) &\to \delta(\mathbf{r}^N - \mathbf{r}'^N) \qquad \text{as} \quad \beta_2 \to \beta_1 \\ &= 0 \qquad \text{if} \quad \beta_2 < \beta_1. \end{aligned} \tag{4.7}$$

As is well known, the kernel may be expanded in terms of the eigenfunctions of the integral equation. From Eq. (4.6), we have

$$\begin{aligned} K(2, 1) &= \sum_n \frac{\psi_n(\mathbf{r}^N)\psi_n^*(\mathbf{r}'^N)}{\lambda_n} \\ &= \sum_n \{\exp[-(\beta - \beta')E_n]\}\psi_n(\mathbf{r}^N)\psi_n^*(\mathbf{r}'^N). \end{aligned} \tag{4.8}$$

In view of this form and Eq. (4.2), we shall consider that the conjugate complex to $\Psi_n(\mathbf{r}^N, \beta)$ is $\Psi_n^*(\mathbf{r}^N, \beta) = \psi_n^*(\mathbf{r}^N)e^{\beta H}$. Accordingly, $\Psi_n(\mathbf{r}^N, \beta)$ is normalized in the same way as $\psi_n(\mathbf{r}^N)$.

Equating β' to 0 in Eq. (4.8), we find that $K(\mathbf{r}^N\beta; \mathbf{r}'^N 0)$ reduces to the density matrix. We arrive at an important relation:

$$Z_N = \int K(\mathbf{r}^N\beta; \mathbf{r}^N 0)\, d\mathbf{r}^N. \tag{4.9}$$

We now investigate the case where the interaction Φ is small compared with H_0. Corresponding to Eq. (4.1), we put

$$K(2, 1) = K_0(2, 1) + K'(2, 1), \tag{4.10}$$

where $K_0(2, 1)$ satisfies

$$\left[\frac{\partial}{\delta\beta_2} + H_0(2)\right]K_0(2, 1) = \delta(2 - 1). \tag{4.11}$$

Using Eqs. (4.10) and (4.11) in Eq. (4.5), we get

$$\frac{\partial K'}{\partial \beta_2} + H_0 K' = -\Phi K. \tag{4.12}$$

However, since the Green's function belonging to this equation is $K_0(2, 1)$, we obtain

$$K'(2, 1) = -\int_{\beta_1}^{\beta_2} \int_V K_0(2, 3)\Phi(3)K(3, 1)\, d3, \tag{4.13}$$

where $d3 = d\mathbf{r}_3\, d\beta_3$. Note that the domain of the β-integration is from β_1 to β_2 and that of the $\mathbf{r}$-integration is for the entire volume. The limitation of the β-domain comes from the discontinuous property of $K_0(2, 3)$ for $(\beta_2 - \beta_1) \to 0$.

Adding $K_0(2, 1)$ to the above equation, we find an integral equation

$$K(2, 1) = K_0(2, 1) - \int K_0(2, 3)\Phi(3)K(3, 1)\, d3. \tag{4.14}$$

This integral equation can be solved by iteration. We arrive at

$$K(2,1) = K_0(2,1) - \int_{\beta_1}^{\beta_2}\int_V K_0(2,3)\Phi(3)K_0(3,1)\,d3 + \int_{\beta_1}^{\beta_2}\int_V\int_{\beta_1}^{\beta_3}\int_V K_0(2,3)\Phi(3)K_0(3,4)\Phi(4)K_0(4,1)\,d3\,d4\cdots. \quad (4.15)$$

The zeroth-order propagator is obtained explicitly by using Eq. (4.8) for ψ_0. For Boltzmann statistics, it is

$$K_0(\mathbf{r}\beta;\mathbf{r}'\beta') = \frac{1}{N!}\prod_{i=1}^{N} K_0(\mathbf{r}_i\beta;\mathbf{r}_i'\beta') = \frac{1}{N!}\prod_{i=1}^{N}\frac{\exp[-m(\mathbf{r}_i-\mathbf{r}_i')^2/\hbar^2 2(\beta-\beta')]}{[(2\hbar^2\pi/m)(\beta-\beta')]^{3/2}}. \quad (4.16)$$

In particular,

$$K_0(\mathbf{r}^N\beta;\mathbf{r}^N 0) = \frac{[(2\pi mkT/h^2)^{3/2}]^N}{N!}. \quad (4.17)$$

On the other hand, if Fermi or Bose statistics have to be used, $K_0(\mathbf{r}\beta;\mathbf{r}'\beta')$ is given by

$$K_0(\mathbf{r}^N\beta;\mathbf{r}'^N\beta') = \frac{1}{N!}\begin{vmatrix} K_0(\mathbf{r}_1\beta;\mathbf{r}_1'\beta'),\ldots,K_0(\mathbf{r}_1\beta;\mathbf{r}_N'\beta') \\ \vdots \qquad\qquad \vdots \\ K_0(\mathbf{r}_N\beta;\mathbf{r}_1'\beta'),\ldots,K_0(\mathbf{r}_N\beta;\mathbf{r}_N'\beta') \end{vmatrix}_{\pm} = \frac{1}{N!}\det{}_N(\mathbf{r}^N\beta;\mathbf{r}'^N\beta'), \quad (4.18)$$

where the plus and minus signs indicate permanent for bosons and determinant for fermions. This expression must be compared with Eq. (3.3).

The partition function of free particles can be obtained by integrating $K_0(\mathbf{r}^N\beta;\mathbf{r}^N 0)$ over the coordinates [see Eq. (4.9)]. If the permanent or determinant on the right-hand side of Eq. (4.18) is expanded, various products of K_0-functions appear, which can be classified by the number of cycles or exchanges of particles. For instance, a typical l-cycle term in the partition function is given by the integral

$$I_l = \int K_0(\mathbf{r}_1\beta;\mathbf{r}_l 0)K_0(\mathbf{r}_l\beta;\mathbf{r}_{l-1}0)\cdots \times K_0(\mathbf{r}_2\beta;\mathbf{r}_1\beta)\,d\mathbf{r}_1\cdots d\mathbf{r}_l. \quad (4.19)$$

In general, if m_l l-cycle terms appear in the expansion, the partition function is expressed as follows:

$$Z_N{}^0 = \frac{1}{N!}N!\sum\prod\frac{1}{m_l!(l!)^{m_l}}[(\pm)^{l-1}(l-1)!\,I_l]^{m_l}, \quad (4.20)$$

where the summation and product are over all l and m_l subject to

$$\sum lm_l = N.$$

$N!/\prod[m_l!\,(l!)^{m_l}]$ is the combinatorial factor to distribute N objects into m_l identical l subgroups; $(l-1)!$ gives the weight of l-cycle diagrams. The factors $(l-1)!$ in the numerator and $l!$ in the denominator together give $(l-1)!/l!$, which is the weight of a ring of l particles when a pair of mirror-image diagrams such as

1 1

2 3 3 2

is counted separately. The factors $(\pm)^{l-1}$ are due to the particle exchanges.

It is convenient to introduce $b_l{}^0$ by

$$b_l{}^0 = \frac{(\pm)^{l-1}}{l}I_l. \quad (4.21)$$

Let z be the absolute activity. The grand partition function becomes

$$\Xi_0 = \sum_N z^N\sum_l\prod_{m_l}\frac{b_l{}^{0m_l}}{m_l!} = \exp\left(\sum_l z^l b_l{}^0\right), \quad (4.22)$$

and thus

$$\frac{pV}{kT} = \sum_l z^l b_l{}^0 \quad (4.23)$$

which is identical in form with Eq. (2.17) of Chapter 5. Note the difference also.

In order to evaluate $b_l{}^0$ explicitly, it is convenient to use the Fourier transform of the free-particle propagator. The single free-particle propagator is

$$K_0(\mathbf{r}\beta;\mathbf{r}'\beta') = \frac{1}{h^3}\int\exp\frac{-p^2(\beta-\beta')}{2m}\exp[-i\hbar^{-1}\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')]\,d\mathbf{p}. \quad (4.24)$$

In other words, $K_0(\mathbf{r}\beta;\mathbf{r}'\beta')$ is obtained by taking the Fourier transform of $K_0(\mathbf{p};\beta-\beta')$, defined by

$$K_0(\mathbf{p};\beta-\beta') = \frac{1}{h^3}\exp\frac{-p^2(\beta-\beta')}{2m}. \quad (4.25)$$

$b_l{}^0$ is easily obtained by a convolution (Faltung) process, leading to

$$b_l{}^0 = \frac{(\pm)^{l-1}(l-1)!\,V}{l!\,h^3}\int_0^\infty 4\pi p^2\exp\frac{-l\beta p^2}{2m}\,dp, \quad (4.26)$$

where $(l-1)!$ is the number of ways $b_l{}^0$ is constructed. Thus,

$$\log \Xi_0 = \mp \frac{4\pi V}{h^3} \int p^2 \log\left(1 \mp z \exp \frac{-\beta p^2}{2m}\right) dp, \tag{4.27}$$

in accord with Eq. (3.13) of Chapter 4. In Eq. (4.27), the upper and lower signs correspond, respectively, to Bose and Fermi statistics.

Problem

Show that

$$K(\mathbf{r}_2\beta_2; \mathbf{r}_1\beta_1) = \int K(\mathbf{r}_2\beta_2; \mathbf{r}_3\beta_3)K(\mathbf{r}_3\beta_3; \mathbf{r}_1\beta_1)\, d\mathbf{r}_3, \tag{4.28}$$

$$\int K^*(\mathbf{r}_2\beta_2; \mathbf{r}_1\beta_1)K(\mathbf{r}_2\beta_2; \mathbf{r}_1'\beta_1)\, d\mathbf{r}_2 = \delta(\mathbf{r}_1 - \mathbf{r}_1'). \tag{4.29}$$

10.5. Graphic Representation of Propagators

In order to treat the general case with interactions, it is convenient to illustrate the terms on the right-hand side of Eq. (4.15) by diagrams. As an example, let us consider the case of three fermions. The expansion of the determinant $\det_3(\mathbf{r}^3\beta; \mathbf{r}^3 0)$ gives the following terms:

$$\begin{aligned}\det_3(\mathbf{r}^3\beta; \mathbf{r}^3 0) = {} & K_0(\mathbf{r}_1\beta; \mathbf{r}_1 0)K_0(\mathbf{r}_2\beta; \mathbf{r}_2 0)K_0(\mathbf{r}_3\beta; \mathbf{r}_3 0) \\ & - [K_0(\mathbf{r}_1\beta; \mathbf{r}_1 0)K_0(\mathbf{r}_2\beta; \mathbf{r}_3 0)K_0(\mathbf{r}_3\beta; \mathbf{r}_2 0) \\ & \quad + K_0(\mathbf{r}_1\beta; \mathbf{r}_3 0)K_0(\mathbf{r}_2\beta; \mathbf{r}_2 0)K_0(\mathbf{r}_3\beta; \mathbf{r}_1 0) \\ & \quad + K_0(\mathbf{r}_1\beta; \mathbf{r}_2 0)K_0(\mathbf{r}_2\beta; \mathbf{r}_1 0)K_0(\mathbf{r}_3\beta; \mathbf{r}_3 0)] \\ & + [K_0(\mathbf{r}_1\beta; \mathbf{r}_3 0)K_0(\mathbf{r}_2\beta; \mathbf{r}_1 0)K_0(\mathbf{r}_3\beta; \mathbf{r}_2 0) \\ & \quad + K_0(\mathbf{r}_3\beta; \mathbf{r}_1 0)K_0(\mathbf{r}_1\beta; \mathbf{r}_2 0)K_0(\mathbf{r}_2\beta; \mathbf{r}_3 0)], \end{aligned} \tag{5.1}$$

where the first term is the diagonal term. The three terms in the first bracket are characterized by one exchange of particles and those in the last bracket by two exchanges.

The propagator $K_0(\mathbf{r}\beta; \mathbf{r}0)$ may be illustrated by a line segment of length β, which starts at the point $(\mathbf{r}, 0)$ and ends at $(\mathbf{r}, \beta)$. It is convenient to connect the end $(\mathbf{r}, 0)$ with $(\mathbf{r}, \beta)$ and make a circular loop of length β. This process corresponds to drawing the diagram on a torus of circumference β. Since a position vector $\mathbf{r}$ corresponds to a particle, one circular loop represents a particle. Therefore, the first term in the above expression may be illustrated by three independent circles such as shown in Fig. 10.1a.

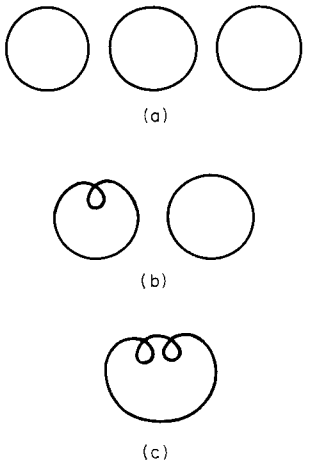


Fig. 10.1. Toron graphs for three free particles.

The terms with one exchange may be illustrated as in Fig. 10.1b, which is composed of a two-loop diagram and one circle. Here, for instance, the particle at $(\mathbf{r}_1, 0)$ is propagated first to $(\mathbf{r}_2, \beta)$ by $K_0(\mathbf{r}_2\beta; \mathbf{r}_1 0)$. Note, $(\mathbf{r}_2, \beta)$ and $(\mathbf{r}_2, 0)$ represent the same point. Then, from $(\mathbf{r}_2, 0)$, the particle propagates to $(\mathbf{r}_1, \beta)$ by $K_0(\mathbf{r}_1\beta; \mathbf{r}_2 0)$. That is, it makes one more loop to go back to the starting point. Thus, the propagator of two particles with one exchange is illustrated by a two-loop diagram. Particle 3 runs simply from $(\mathbf{r}_3, 0)$ to $(\mathbf{r}_3, \beta)$ and is represented by a circle.

It is interesting to observe the indistinguishability of particles in the two-loop diagram. This diagram is due to two particles, but since the curve is continuous, we cannot tell which part of the curve corresponds to which particle. The two loops simply represent two particles with one exchange. We may say that these particles are bound together by a statistical pseudo-force. The diagram may then be considered to represent a hypothetical "molecule" composed of two particles which may be called a "2-*toron*." Since one exchange is involved, this type of diagram must have a plus or a minus sign depending on the statistics of the particles.

The terms in the last bracket in Eq. (5.1) are characterized by two exchanges between the three particles. We illustrate these terms by two-loop diagrams as in Fig. 10.1c. In this diagram, we have the propagation

$$(\mathbf{r}_1, 0) \to (\mathbf{r}_2, \beta) = (\mathbf{r}_2, 0) \to (\mathbf{r}_3, \beta) = (\mathbf{r}_3, 0) \to (\mathbf{r}_1, \beta).$$

Again we have a continuous curve showing the quantum mechanical indistinguishability of the particles. Due to two exchanges, this type of graph always has a plus sign irrespective of the statistics. We say that such a diagram represents a 3-toron.

The above consideration of $\det_3(\mathbf{r}^3\beta;\mathbf{r}^30)$ can be easily generalized to $\det_N(\mathbf{r}^N\beta;\mathbf{r}^N0)$. If the expansion of $\det_N(\mathbf{r}^N\beta;\mathbf{r}^N0)$ gives m_s torons composed of s particles, we have, clearly,

$$\sum sm_s = N. \tag{5.2}$$

Since there is no interaction between the torons, the total system may be considered equivalent to an ideal gas mixture of these toron molecules. Therefore, we expect the equation of state to be in the form

$$\frac{pV}{kT} = \sum_s m_s, \tag{5.3}$$

as in the case of Eq. (2.23) of Chapter 5 for an imperfect gas.

The term $m_s = b_s{}^0 z^s$ represents the statistical abundance of s-torons and can be given explicitly by statistical considerations of the relative probability of torons in an analogous way as in Section 5.2. Thus, we start with forming an s-toron by collecting s particles. The probability of finding an s-toron relative to that for finding a 1-toron (namely, with one particle) may be obtained in the following way. An s-toron is composed of s particles with the total energy $\sum_{i=1}^{s}(p_i{}^2/2m)$, which may be written as

$$\frac{sp^2}{2m} = \sum_{i=1}^{s} \frac{p_i{}^2}{2m}, \tag{5.4}$$

where $p^2/2m$ is the average kinetic energy per particle of the s-toron. The term $sp^2/2m$ may be interpreted as the energy of an s-toron. Then,

$$\exp\left(-\frac{\beta sp^2}{2m}\right)\frac{d\mathbf{p}}{h^3}$$

is the probability of an s-toron being in the energy state specified by $\mathbf{p}$. The total relative probability $b_s{}^0$ is obtained from

$$b_s{}^0 = (\pm)^{s-1}\frac{(s-1)!}{s!}V\int \exp\left(-\frac{\beta sp^2}{2m}\right)\frac{d\mathbf{p}}{h^3}, \tag{5.5}$$

where $(s-1)!$ is the weight due to the permutation of s particles forming an s-toron and $1/(s!)$ is the factor due to the indistinguishability of particles. In the cases of bosons and fermions, permutations yield a plus or minus sign depending on even or odd permutations. This sign factor should be included in b_s.

Finally, combination of Eqs. (5.3), (5.4), and (5.5) gives

$$\frac{p}{kT} = \mp\frac{4\pi}{h^3}\int p^2 \log\left(1 \mp z\exp\frac{-p^2}{2mkT}\right)dp,$$

which coincides with Eq. (4.27).

10.6. Linked Cluster Expansion of the Equation of State

The partition function of an interacting electron gas is obtained by taking into account the trace of the iteration series of Eq. (4.15), which is expressed in ascending powers of the potentials. We shall rearrange the series and introduce the grand partition function to obtain a cluster series.

The contribution to Z_N from connected graphs of l particles is denoted by $l!\,b_l$:

$$b_l = \frac{1}{l!}\sum (\text{integral associated with a diagram of } l \text{ connected particles}) \times (\text{number of ways the diagram is constructed}), \tag{6.1}$$

where the summation is over all possible connected diagrams of l particles.

Let us assume that our system has N electrons in volume V. From quantum statistics, the electrons form various torons corresponding to different ways of subgrouping the natural number N:

$$\sum_l lr_l = N. \tag{6.2}$$

Here, r_l is the number of l-torons. As we have seen before, the number of ways of dividing an integer N into r_l subgroups of l members is given by $N!/\prod[(l!)^{r_l}r_l!]$. Combining this with the factor $1/N!$ appearing in the definition of the propagator, we find that the clusters of electrons have the weight

$$\frac{1}{N!}N!\frac{1}{\prod(l!)^{r_l}r_l!} = \frac{1}{\prod(l!)^{r_l}r_l!}. \tag{6.3}$$

By multiplying this by $(l!\,b_l)^{r_l}$, the partition function can be written as

$$Z_N = \sum_{\sum lr_l = N}\prod\frac{b_l^{r_l}}{r_l!}, \tag{6.4}$$

where the summation and product are for all integers l and r_l satisfying the condition

$$\sum lr_l = N.$$

The grand partition function is obtained from Eq. (6.4) as follows:

$$\Xi = \sum_N z^N Z_N = \sum_N\left[\sum\prod\frac{(z^lb_l)^{r_l}}{r_l!}\right] = \exp\left(\sum_l z^l b_l\right). \tag{6.5}$$

Thus, we have arrived at a cluster expansion formula:

$$\log \Xi = \sum b_l z^l. \tag{6.6}$$

The cluster integrals b_l are obtained by analyzing connected diagrams of l particles, some of which may be forming torons. Therefore, let us consider subgrouping l particles into r_s s-subgroups and connecting the subgroups by n interaction lines occurring at various intermediate reciprocal temperatures. The appropriate weight factor to b_l is

$$\frac{1}{l!}\frac{l!}{\prod_s (s!)^{r_s} r_s!}\prod_s r_s!\,[(s-1)!]^{r_s}\frac{(\sum_s r_s)!}{S}(-1)^n\prod(-1)^{(s-1)r_s}$$

$$= (-1)^{n+l-\sum_s r_s}\prod_s \frac{1}{s^{r_s}}\frac{(\sum_s r_s)!}{S}, \tag{6.7}$$

where the first factor $1/l!$ comes from the definition of b_l, the second is from the number of ways of subgrouping l into r_s s-members, and the fourth factor $(s-1)!$ is the weight of an s-toron because it is an ordered ring of s particles.[1] The fifth factor, $(\sum_s r_s)!/S$, comes from the interchange of torons. Here, S is a symmetry factor to be determined by the shape of a graph. For instance, if $(\sum_s r_s)$ torons form a ring, then $S = 2\sum_s r_s$, and the interchange weight is

$$\frac{1}{2}\left(\sum_s r_s - 1\right)!.$$

Finally, the sign factor to b_l is determined by the numbers of interaction lines and of torons. If r_s s-torons are connected by n interaction lines, the sign is

$$(-1)^n\prod(-1)^{(s-1)r_s} = (-1)^{n+l-\sum_s r_s}.$$

In connection with the weight factor, an s-toron, being formed originally by s particles, may have interaction lines in s positions. That is, the contribution b_l from an s-toron in a connected l-particle diagram must include all positions of the interactions in the toron. Moreover, an interaction line can occur at various reciprocal temperatures. If a graph has n interaction lines occurring at n different reciprocal temperatures arranged such that

$$\beta^{(n)} > \beta^{(n-1)} > \cdots > \beta'' > \beta', \tag{6.8}$$

[1] An s toron has s positions for an interaction line, and $(s-1)!$ must be multiplied by s.

similar graphs may appear in $n!$ different ways by interchanging the interactions. However, the graphs having two or more interaction lines at the same point (namely, the same temperature) should be eliminated from consideration because they lead to a divergence, as can be seen easily. Because of this, all the β's in the above inequality are assumed different, resulting in a factor $n!$.

Figure 10.2 illustrates first few graphs for b_2 and b_3. Note that some of the

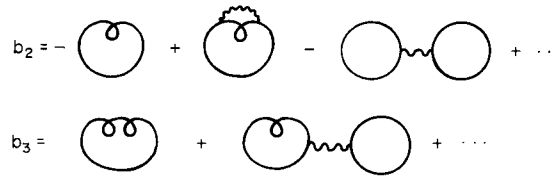


Fig. 10.2. Toron graphs for a cluster series.

graphs result in a divergence and should actually be eliminated. These are the so-called self-energy-type diagrams, where the torons absorb and emit the same interactions within 1-toron propagation.

Due to the combinatorial factors and multiple integrals, the calculation of b_l is not a simple matter. Moreover, it is in general not possible to sum over all the diagrams. Nevertheless, there are many cases where one can treat an infinite number of diagrams of a particular type and obtain important information concerning many-particle systems.

10.7. Equation of State in the Ring Diagram Approximation

Let us consider the second-order perturbation term corresponding to the process illustrated in Fig. 10.3. Here, a particle at $(\mathbf{r}, 0)$ in the coordinate-

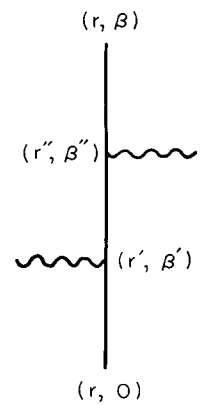


Fig. 10.3. A propagation with two interaction lines.

reciprocal-temperature space propagates to $(\mathbf{r}, \beta)$ after interacting with some other particles of coordinates $(\mathbf{r}', \beta')$ and $(\mathbf{r}'', \beta'')$, the interactions taking place at β' and β''. According to Eq. (4.15), the contribution to the partition function from this process is

$$z \int d\mathbf{r}\, d\mathbf{r}'\, d\mathbf{r}'' \int_0^{\beta} d\beta'' \int_0^{\beta''} d\beta' K_0(\mathbf{r}\beta; \mathbf{r}''\beta'')\phi(\mathbf{r}_l'' - \mathbf{r}'')K_0(\mathbf{r}''\beta''; \mathbf{r}'\beta') \times \phi(\mathbf{r}_2'' - \mathbf{r}')\, K_0(\mathbf{r}'\beta'; \mathbf{r}, 0). \tag{7.1}$$

The Coulomb potential $\phi(r) = e^2/r$ may be Fourier-transformed as follows:

$$\phi(r) = \frac{1}{h^3} \int u(q) \exp(-i\hbar^{-1}\mathbf{q} \cdot \mathbf{r})\, d\mathbf{q},$$

$$u(q) = \int \phi(r) \exp(i\hbar^{-1}\mathbf{q} \cdot \mathbf{r})\, d\mathbf{r}$$

$$= \hbar^2 \left(\frac{4\pi e^2}{q^2}\right). \tag{7.2}$$

This transformation is achieved by introducing a convergence factor $e^{-\alpha r}$ to the Coulomb potential and, after transforming $e^{-\alpha r}/r$, bringing α to zero.

The propagators in the integrand of Eq. (7.1) may be expressed in momentum space using expressions such as Eq. (4.25). If, as in Fig. 10.3, $\mathbf{p}$, $\mathbf{p}'$, and $\mathbf{p}''$ are the momenta representing the free-particle propagations, then

$$K_0(\mathbf{r}'\beta'; \mathbf{r}0) = \frac{1}{h^3} \int \exp\left(\frac{-\beta' p^2}{2m}\right) \exp[-i\hbar^{-1}\mathbf{p} \cdot (\mathbf{r}' - \mathbf{r})]\, d\mathbf{p}, \quad \text{etc.} \tag{7.3}$$

Note here that the coordinates of a particle appear only in the exponential factor. In Eq. (7.1), the coordinates $\mathbf{r}'$ appear only in $\exp[-i\hbar^{-1}(\mathbf{p} - \mathbf{p}' - \mathbf{q}) \cdot \mathbf{r}']$. Thus, the integration of Eq. (7.1) over $\mathbf{r}'$ gives a δ-function:

$$\frac{1}{h^3} \int \exp[-i\hbar^{-1}(\mathbf{p} - \mathbf{p}' - \mathbf{q}) \cdot \mathbf{r}']\, d\mathbf{r}' = \delta(\mathbf{p} - \mathbf{p}' - \mathbf{q}). \tag{7.4}$$

This indicates momentum conservation:

$$\mathbf{p}' = \mathbf{p} - \mathbf{q}. \tag{7.5}$$

Since $\mathbf{p}$ and $\mathbf{p}'$ correspond respectively to the momenta before and after the interaction, we shall say that the particle lost momentum $\mathbf{q}$ (the interaction left the graph). Had we used $-\mathbf{q}$ instead of $\mathbf{q}$, we would have obtained $\mathbf{p}' = \mathbf{p} + \mathbf{q}$. In this case, we would say that the particle gained momentum $\mathbf{q}$ (the interaction entered the graph).

Similar momentum conservations follow from the integrations over $\mathbf{r}''$ and $\mathbf{r}$ of Eq. (7.1):

$$\mathbf{p}'' = \mathbf{p}' - \mathbf{q}',$$
$$\mathbf{p}'' = \mathbf{p}. \tag{7.6}$$

where $\mathbf{q}'$ is the momentum representing the interaction $\phi(\mathbf{r}_l'' - \mathbf{r}'')$. Combining Eqs. (7.5) and (7.6), we find

$$\mathbf{q}' = -\mathbf{q}. \tag{7.7}$$

In other words, the momentum that enters an unlabeled 1-toron through an interaction must go out again. The momentum conservation laws is used frequently in the calculation of b_l. As a result, the factors in Eq. (7.1) become

$$\frac{z}{h^6} \int \exp\left(-\frac{(\beta - \beta'')p^2}{2m}\right) \exp\left(-\frac{(\beta'' - \beta')(\mathbf{p} - \mathbf{q})^2}{2m}\right) \exp\left(-\frac{\beta' p^2}{2m}\right) \times u^2(q) \exp[-i\hbar^{-1}\mathbf{q} \cdot (\mathbf{r}_l'' - \mathbf{r}_2'')]\, d\mathbf{p}\, d\mathbf{q}. \tag{7.8}$$

The exponential factors are simple, and if we use the momentum conservation law from the beginning, we can obtain these factors directly from the diagram. The usefulness of diagrams lies in this simplicity. In (7.8), the first three energy terms on the exponential factors correspond, respectively, to a particle of momentum $\mathbf{p}$ existing between 0 and β', a particle of momentum $\mathbf{p} - \mathbf{q}$ in the interval $\beta'' - \beta'$, and a particle of momentum $\mathbf{p}$ in an interval $\beta - \beta''$. An exiting interaction has a plus sign in the phase factor while an entering interaction has a minus sign. The $\mathbf{p}$-integration of Eq. (7.8) can be carried out, giving

$$\frac{1}{h^3} \int_0^{\beta} d\beta'' \int_0^{\beta''} d\beta' G_1(\mathbf{q}; \beta'', \beta')u^2(q) \exp[-i\hbar^{-1}\mathbf{q} \cdot (\mathbf{r}_l'' - \mathbf{r}_2'')]\, d\mathbf{q}, \tag{7.9}$$

where, with λ for the thermal de Broglie wavelength and $\alpha = |\beta'' - \beta'|$,

$$G_1(\mathbf{q}; \beta'', \beta') = \frac{z}{h^3} \int \exp\left[-\frac{(\beta - \alpha)p^2}{2m} - \frac{\alpha(\mathbf{p} - \mathbf{q})^2}{2m}\right] d\mathbf{p}$$

$$= z\lambda^{-3} \exp\left[-\frac{\alpha(\beta - \alpha)q^2}{2m\beta}\right]. \tag{7.10}$$

Dropping one $u(q)$ from Eq. (7.9), we introduce a new propagator for a particle with one interaction:

$$F(\mathbf{r}_2'' - \mathbf{r}_l''; \beta'' - \beta') = \frac{1}{h^3} \int G_1(\mathbf{q}; \beta'' - \beta')u(q) \exp[i\hbar^{-1}\mathbf{q} \cdot (\mathbf{r}_l'' - \mathbf{r}_2'')]\, d\mathbf{q}. \tag{7.11}$$

This represents a unit process involved in a ring diagram. In this expression, only the difference of the coordinates $\mathbf{r}_2'' - \mathbf{r}_l''$ appears in the exponential function. Thus, the contribution to b_l from l 1-torons forming a ring is given by

$$l!\, b_l = (-1)^l l!\, \tfrac{1}{2}[(l-1)!] \int_0^\beta d\beta^{(l)} \int_0^{\beta^{(l)}} \cdots \int_0^{\beta''} d\beta'$$

$$\times \int \cdots \int F(\mathbf{r}_2' - \mathbf{r}_3'') F(\mathbf{r}_3'' - \mathbf{r}_4''') \cdots F(\mathbf{r}_1^{(l)} - \mathbf{r}_2')\, d\mathbf{r}_2'\, d\mathbf{r}_3'' \cdots d\mathbf{r}_1^{(l)}, \quad (7.12)$$

where for simplicity the β-coordinates in the arguments of the F's are omitted; $l!$ comes from the interchange of l 1-torons; and $(l-1)!/2$ is the weight of an l-ring.

Using Eq. (7.11) and carrying out the integrations, we find that all the interactions are represented by the same $\mathbf{q}$, as a result of the coordinate integrals. We also note that G_1 depends only upon $\alpha = |\beta'' - \beta'|$. Thus, all the β-integrations can be carried over the range between 0 and β provided we introduce a factor $1/l!$. We obtain

$$b_l z^l = \frac{(-1)^l (l-1)!}{2l!\, h^3} \int_0^\beta \cdots \int_0^\beta d\beta'' \int_0^\beta d\beta' u^l(q) G_1(\mathbf{q}; \beta', \beta'')$$

$$\times G_1(\mathbf{q}; \beta'', \beta''') \cdots G_1(\mathbf{q}; \beta^{(l)}, \beta')\, d\mathbf{q}. \quad (7.13)$$

The multiple integrations on the right-hand side can be performed in terms of the eigenvalues of the integral equation:

$$\lambda\psi(\beta') = \int_0^\beta G_1(\mathbf{q}, |\beta'' - \beta'|)\psi(\beta'')\, d\beta''. \quad (7.14)$$

The kernel is expanded by using the normalized eigenfunctions as follows:

$$G_1(\mathbf{q}, |\beta'' - \beta'|) = \sum \lambda_k \psi_k(\beta') \psi_k^*(\beta''). \quad (7.15)$$

Due to the periodic character of torons, we require

$$G_1(\mathbf{q}, \alpha) = G_1(\mathbf{q}, \alpha + \beta). \quad (7.16)$$

Then, the normalized characteristic function is

$$\psi_j(\beta') = \beta^{-1/2} \exp\left(\frac{-2\pi i j \beta'}{\beta}\right), \quad (7.17)$$

and the eigenvalues are

$$\lambda_j = \frac{z}{\lambda^3} \int_0^\beta \exp\left[-\frac{q^2\alpha(\beta - \alpha)}{2m\beta} - \frac{2\pi i j \alpha}{\beta}\right] d\alpha. \quad (7.18)$$

Carrying out the multiple integrations of Eq. (7.13), we find

$$b_l z^l = \frac{V}{2lh^3} \int \sum_j [-u_j(q)\lambda_j(q)]^l\, d\mathbf{q}. \quad (7.19)$$

In order to obtain the grand partition function, it is necessary to sum over all l. We arrive at the formula

$$\log \frac{\Xi}{\Xi_0} = \sum_{l=2}^{\infty} z^l b_l$$

$$= \frac{V}{2h^3} \int \sum_j \{u(q)\lambda_j - \log[1 + u(q)\lambda_j]\}\, d\mathbf{q}, \quad (7.20)$$

where Ξ_0 is the grand partition function of an ideal gas, given by Eq. (4.27). This result was first obtained by Gell-Mann and Brueckner (*12*) (for energy) and then by Montroll and Ward (*13*) in the above form.

10.8.* The Eigenvalues of Quantum Propagators

Equation (7.20) holds for fermions and bosons forming a ring if the kernel function G is properly defined for the elements of a ring. In this quantal case, we meet the following situation:

(*i*) *Formation of a ring*

An l-toron ring is formed by r_t t-torons, with integers r_t and t satisfying $\sum t r_t = l$.

(*ii*) *Momentum conservation*

Integrations over the coordinates of the particles forming a ring result in momentum conservation: All the interactions are represented by the same momentum $\mathbf{q}$.

(*iii*) *Weight factor*

According to Eq. (6.7), the weight factor is

$$(-1)^l \prod \frac{1}{t^{r_t}} (\textstyle\sum r_t)! \, \frac{(\sum r_t - 1)!}{2},$$

where $\sum r_t$ represents the total number of interaction lines or torons.

(*iv*) *Definition of the G-function*

The contribution of each t-toron in a ring will be defined in such a way that $1/t$ in the above weight factor is included in the contribution. One of the

two interaction lines in a toron may be originated from any particle forming the toron. Thus, the given t-toron gets a factor t from this possibility, which cancels $1/t$ in the above weight factor. Another interaction line can also be at various positions, and the contribution to b_l from the t-torons is obtained by summing over all these possibilities.

(*v*) *Enlargement of the integration domain*

In the above weight factor, $(\sum r_t)!$ may be dropped out by removing the restrictions on $(\sum r_t)$ β-integrations [see Eq. (7.13)].

(*vi*) *Formation of effective elements*

In the grand ensemble, each ring element will be replaced by an effective element that is the sum of torons of all orders (Fig. 10.4). If all the ring elements are identical, another factor $1/(\sum r_t)!$ must be introduced.

(*vii*) *Summation over all rings*

If l is arbitrary, $\sum r_t$ is also arbitrary. Final grand ensemble sum over all $\sum r_t$ yields Eq. (7.20) provided that the λ_j are defined properly. Since $(-1)^l$ in Eq. (7.13) is due to interactions, each toron must be given a proper sign.

The G-function for this case is obtained by replacing each component toron in a ring by an effective element with two interaction lines, the effective element being the sum of torons of all orders as in Fig. 10.4.

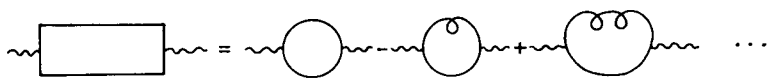


Fig. 10.4. An effective toron diagram.

Let us consider an l-toron in a ring of fermions. Measuring from an arbitrary chosen point of the l-toron, we shall assume that an absorption interaction line occurs at the reciprocal temperature β' and an emission line at $j\beta + \beta''$, where $\beta'' > \beta'$, and j is an arbitrary integer between 0 and $l-1$. The corresponding propagator is

$$u(q)\frac{1}{h^3}\sum_{j=0}^{l-1}\int \exp\left\{-\frac{[(l-j)\beta-\alpha]p^2}{2m}-\frac{(j\beta+\alpha)(\mathbf{p}+\mathbf{q})^2}{2m}\right\}d\mathbf{p}\,.$$

Here, if we use $\alpha = |\beta'' - \beta'|$, the same expression holds even for $\beta'' < \beta'$. The kernel corresponding to the effective element of Fig. 10.4 is obtained by multiplying the above expression by $(-1)^{l+1}z^l$ and summing over all l. Thus, we have a quantal propagator for an effective element of a ring:

$$\begin{aligned}G(\mathbf{q},\alpha) &= \sum_{l=1}^{\infty}(-1)^{l+1}z^l\sum_{j=0}^{l-1}\frac{1}{h^3}\int d\mathbf{p}\\ &\quad\times\exp\left\{-\frac{[(l-j)\beta-\alpha]p^2}{2m}-\frac{(j\beta+\alpha)(\mathbf{p}+\mathbf{q})^2}{2m}\right\}\\ &= \frac{1}{h^3}\int\frac{z\exp(-\beta p^2/2m)}{1+z\exp(-\beta p^2/2m)}\,\frac{1}{1+z\exp[-\beta(\mathbf{p}+\mathbf{q})^2/2m]}\\ &\quad\times\exp\left\{-\frac{[(\mathbf{p}+\mathbf{q})^2-\mathbf{p}^2]\alpha}{2m}\right\}d\mathbf{p}.\end{aligned}\tag{8.1}$$

The eigenvalues of this kernel can be obtained as follows:

$$\begin{aligned}\lambda_j(q) &= \frac{1}{h^3}\int d\mathbf{p}\,f^{(0)}(p)[1-f^{(0)}(\mathbf{p}+\mathbf{q})]\\ &\quad\times\int_0^{\beta}\exp\frac{\alpha[p^2-(\mathbf{p}+\mathbf{q})^2]}{2m}\exp\frac{2\pi ij\alpha}{\beta}\,d\alpha.\end{aligned}\tag{8.2}$$

We shall find in Section 15.7 that these eigenvalues characterize the dielectric response of a plasma.

The last integral on the right may be transformed, using the identity

$$\frac{1}{2\pi}\int\frac{\exp ias}{is+(p^2/2m)}\,ds=\begin{cases}\exp\left(-\dfrac{ap^2}{2m}\right), & a>0,\\ 0, & a<0,\end{cases}\tag{8.3}$$

into the formula

$$\begin{aligned}\lambda_j(q) &= \frac{1}{(2\pi)h^3}\int d\mathbf{p}\int_{-\infty}^{\infty}ds\,\frac{ze^{is\beta}}{1+ze^{is\beta}}\\ &\quad\times\frac{1}{(p^2/2m)+is}\,\frac{1}{[(\mathbf{p}+\mathbf{q})^2/2m]+i(s-2\pi j\beta^{-1})}.\end{aligned}\tag{8.4}$$

This formula has been derived for fermions. It is clear that a similar formula holds for bosons.

Equation (8.2) or (8.4) yields

$$\begin{aligned}\lambda_j(q) &= \frac{z}{h^3}\int d\mathbf{p}\,\frac{[\exp(-\beta p^2/2m)]-\exp[-\beta(\mathbf{p}+\mathbf{q})^2/2m]}{[(\mathbf{p}+\mathbf{q})^2/2m]-(p^2/2m)-(2\pi ij/\beta)}\\ &\quad\times\frac{1}{1+z\exp(-\beta p^2/2m)}\,\frac{1}{1+z\exp[-\beta(\mathbf{p}+\mathbf{q})^2/2m]},\end{aligned}\tag{8.5}$$

where we notice the appearance of the Fermi distribution function:

$$f^{(0)}(p) = \frac{z \exp(-\beta p^2/2m)}{1 + z \exp(-\beta p^2/2m)}. \tag{8.6}$$

Changing the variable of integration from $\mathbf{p}$ to $-(\mathbf{p}+\mathbf{q})$ in the second term, one finds

$$\begin{aligned}\lambda_j(q) &= \frac{8\pi m}{h^3}\int_0^\infty p^2\, dp\, f^{(0)}(p) \\ &\quad \times \int_0^\pi \sin\theta\, d\theta \frac{q^2 + 2pq\cos\theta}{(q^2 + 2pq\cos\theta)^2 + (4\pi jm/\beta)^2} \\ &= \frac{2\pi m}{h^3 q}\int_{-\infty}^\infty p\, dp\, f^{(0)}(p) \log\left[(q^2+2pq)^2 + \left(\frac{4\pi jm}{\beta}\right)^2\right].\end{aligned} \tag{8.7}$$

10.8.1. Fermion Eigenvalues at Low Temperature

In the limit $\beta \to \infty$, the fermion distribution is

$$f^{(0)}(p) = \begin{cases} 0, & p^2 > p_0^2, \\ 1, & p^2 < p_0^2, \end{cases} \tag{8.8}$$

Thus, from Eq. (8.7), we obtain

$$\begin{aligned}\lambda_j(q) &= \frac{\pi m}{2h^3 q^3}\int_{-2p_0q}^{2p_0q} x\, dx \log\left[(x+q^2)^2 + \left(\frac{4\pi jm}{\beta}\right)^2\right] \\ &= \frac{\pi m}{2h^3 q^3}\int_{q^2-2p_0q}^{q^2+2p_0q} dy(y-q^2)\log\left[y^2 + \left(\frac{4\pi jm}{\beta}\right)^2\right] \\ &= \frac{2m\pi p_0}{h^3} F(s, y),\end{aligned} \tag{8.9}$$

where

$$\begin{aligned}F(s,y) = 1 &- \frac{1}{8}\left(s - \frac{4}{s} - \frac{y^2}{s^3}\right)\log\left|\frac{(s^2+2s)^2 + y^2}{(s^2-2s)^2 + y^2}\right| \\ &- \frac{y}{2s}\left[\tan^{-1}\left(\frac{s^2+2s}{y}\right) - \tan^{-1}\left(\frac{s^2-2s}{y}\right)\right].\end{aligned} \tag{8.10}$$

In these expressions, we have used

$$s = \frac{q}{p_0}, \qquad x = \frac{p_0 r}{\hbar}, \qquad y = \frac{4\pi jm}{\beta p_0^2}, \tag{8.11}$$

$$p_0 = \hbar(3\pi^2 n)^{1/3}. \tag{8.12}$$

Equation (8.10) is illustrated in Fig. 10.5 for several values of y. Note that only the curve corresponding to $y = 0$ is not zero at $s = 0$.

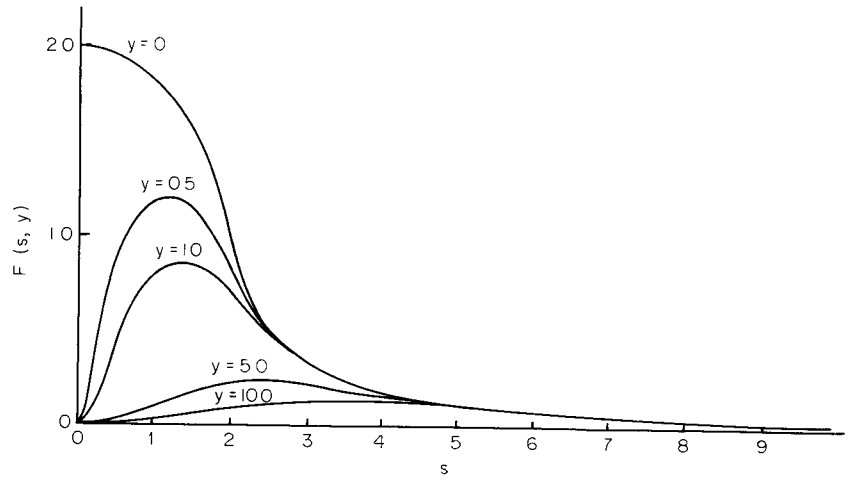


Fig. 10.5. The function $F(s, y)$ for the eigenvalues of degenerate electrons. The same function appears in the dielectric function of an electron gas.

10.8.2. Boson Eigenvalues at Low Temperatures

The eigenvalue expression (8.2) for fermions can be used for bosons if $1 - f^{(0)}$ is replaced by $1 + f^{(0)}$, with $f^{(0)}$ given by

$$f^{(0)}(p) = \frac{z \exp(-\beta p^2/2m)}{1 - z \exp(-p^2/2m)}. \tag{8.13}$$

At low temperatures, the particles of zero momentum may be separated out from the integral for boson eigenvalues. We consider processes in which these particles acquire momenta in the interval $\beta'' - \beta'$ and then drop again into the zero-momentum state. Accordingly, their contribution to the eigenvalues is found as follows:

$$\lambda_j^0(q) = \frac{4mn_0[1 - \exp(-\beta q^2/2m)]}{[1 - z\exp(-\beta q^2/2m)][q^2 + (4\pi jm/\beta q)^2]}, \qquad q \neq 0. \tag{8.14}$$

At very low temperatures, near 0°K, this may be approximated by

$$\lambda_j^0(q) \approx \frac{4mn_0}{q^2 + (4\pi jm/\beta q)^2}. \qquad q \neq 0, \tag{8.15}$$

where n_0 is the number density of the particles in the lowest energy state. Note that this approximation is correct for $z = 1$ in Eq. (8.14). On the other hand, for $j = 0$ and $q \to 0$, Eq. (8.14) approaches a finite value, while Eq. (8.15) does not.

The contribution from the excited particles to the eigenvalues is evaluated as follows:

$$\lambda_j^1(q) = \frac{2m\pi}{h^3 q} \sum_l z^l \int_{-\infty}^{\infty} p\left(\exp - \frac{\beta l p^2}{2m}\right) \log\left[(q^2 + 2pq)^2 + \left(\frac{4\pi jm}{\beta}\right)^2\right] dp$$

$$= \frac{m^2}{\hbar^3 \pi^2 \beta} \sum_l \frac{z^l}{l} \int_{-\infty}^{\infty} \left(\exp - \frac{\beta l p^2}{2m}\right) \frac{q + 2p}{q[(q + 2p)^2 + (4\pi jm/\beta q)^2]} dp$$

$$= 4mn_1 \left[q^2 + \left(\frac{4\pi jm}{\beta q}\right)^2\right]^{-1} \left\{1 - \frac{12m}{\beta q^2[1 + (4\pi jm/\beta q^2)^2]} + \cdots\right\}, \quad (8.16)$$

where n_1 is the number density of the excited particles.

Combining Eqs. (8.15) and (8.16), one finds, in the first approximation and below the λ-point,

$$\lambda_j(q) = \lambda_j^0(q) + \lambda_j^1(q)$$

$$= 4mn\left[q^2 + \left(\frac{4\pi jm}{\beta q}\right)^2\right]^{-1}$$

$$- 48m^2 n_1 \left[q^2 + \left(\frac{4\pi jm}{\beta q}\right)^2\right]^{-1} \left\{\beta q^2 \left[1 + \left(\frac{4\pi jm}{\beta q^2}\right)^2\right]^{-1}\right\}. \quad (8.17)$$

Above the λ-point, λ_j is given by equating n_1 to n in the same expression.

10.9. Correlation Energy of an Electron Gas

As an example of using the formula (7.20), let us evaluate the energy of an electron gas at zero temperature. We remark first that the Hamiltonian can be expressed in the following form:

$$H = \sum_i \frac{p_i^2}{2m} + \sum_{i<j} \frac{e^2}{r_{ij}}$$

$$= \frac{me^4}{\hbar^2 r_s^2} \left[\sum_i \frac{p_i'^2}{2} + r_s \sum_{i<j} \frac{1}{r_{ij}'}\right], \quad (9.1)$$

where

$$r_s = \frac{r_0}{a_0} = \left(\frac{3}{4\pi n}\right)^{1/3} \frac{me^2}{\hbar^2}. \quad (9.2)$$

The quantity r_0 is the radius of the Fermi sphere and $a_0 = \hbar^2/me^2$ is the Bohr radius. In Eq. (9.1), the momentum is measured in the unit $\hbar/r_0$, and the distance by r_0. These units are called the Rydberg units.

Equation (9.1) shows that the kinetic energy is proportional to r_s^{-2}. In rydbergs, the Fermi energy per particle is given by

$$\varepsilon_0 = \frac{\frac{3}{5}(p_0^2/2m)}{(e^4 m/2\hbar^2)} = \frac{3}{5}\left(\frac{9\pi}{4}\right)^{2/3} \frac{1}{r_s^2}$$

$$\sim \frac{2.21}{r_s^2}. \quad (9.3)$$

The value of r_s is 3.22 for lithium, 3.96 for sodium, and 4.87 for potassium. Thus, the Fermi energy is of the order of 0.1.

The values of r_s decrease with increasing density. Thus, r_s may be considered to be a small coupling parameter for an electron gas at high density. The next-order term in the ground-state energy which follows ε_0 is expected to be proportional to r_s^{-1}.

Indeed, the next-order term, known as the *exchange energy*, is given by

$$\varepsilon_x = -\frac{\frac{3}{4}(e^2/\pi)p_0/\hbar}{(e^4 m/2\hbar^2)} = -\frac{3}{2\pi}\left(\frac{9\pi}{4}\right)^{1/2} \frac{1}{r_s}$$

$$\sim -\frac{0.916}{r_s}. \quad (9.4)$$

If the effect of the interactions is evaluated in second-order perturbation, we expect a term which is one order higher than the above exchange energy. Unfortunately, the long-range Coulomb forces cause a logarithmic divergence in the second-order perturbation. Therefore, a more elaborate evaluation is needed for higher-order terms.

The difference from the ground state energy ε_g

$$\varepsilon_c = \varepsilon_g - \varepsilon_0 - \varepsilon_x \quad (9.5)$$

is called the *correlation energy*. It is an important quantity in solid-state physics, and there have been many attempts to evaluate this energy.

A reliable result concerning ε_c has been obtained by Gell-Mann and Brueckner (*12*). Their basic idea is to sum the increasingly divergent perturbation terms before making the final integration over the momentum expressing the Coulomb interaction. Since the Coulomb potential has a momentum dependence $1/q^2$, it is clear that small momentum transfer between the particles is important.

One can imagine that the greatest piling up of factors $1/q^2$ occurs when only a single momentum transfer is involved, and this single q is handed from

electron to electron, contributing a factor $1/q^2$ each time. Effectively, such electron configurations may be represented by ring diagrams. Then we can make use of the formula (7.20). Note that the grand partition function yields the internal energy as follows:

$$\varepsilon = -\frac{1}{N}\frac{\partial(\log \Xi)}{\partial(\log \beta)}. \tag{9.6}$$

For a small momentum transfer, the eigenvalue expression (8.9) may be approximated as follows (*14*, *15*)

$$\lambda_j = \frac{2mp_0}{\hbar^3 8\pi^2} F(s, y),$$
$$F(s, y) = 2\left(1 - \frac{y}{2s}\tan^{-1}\frac{2s}{y}\right). \tag{9.7}$$

Furthermore, the summation in Eq. (7.20) may be replaced by integration:

$$\sum_i \to \frac{\beta {p_0}^2}{4\pi m}\int dy. \tag{9.8}$$

We obtain (13–15)

$$\log\frac{\Xi}{\Xi_0} = \frac{V\beta {p_0}^5}{8\pi m\hbar^3}\iint \{u(q)\lambda(s, y) - \log[1 + u(q)\lambda(s, y)]\}\, ds\, dy. \tag{9.9}$$

Using these results and taking into consideration exchange graphs, Gell–Mann and Brueckner (*12*) found that the correlation energy (in rydbergs) is

$$\varepsilon_c = 0.0622 \log r_s - 0.096 + \cdots. \tag{9.10}$$

The total ground-state energy per particle of an interacting electron gas is given by (*17*)

$$\varepsilon_g = 2.21 r_s^{-2} - 0.916 r_s^{-1} + 0.0622 \log r_s - 0.096 + \cdots. \tag{9.11}$$

Similarly, the specific heat may be expressed in ascending powers of r_s. According to Gell–Mann (*16*) this is given by

$$c_V = {c_V}^0\left\{1 + \frac{r_s}{3\pi^{3/2}}\left[-\log r_s + \log\left(\frac{3\pi^{3/2}}{2}\right) - 2\right] + \cdots\right\}^{-1}, \tag{9.12}$$

where ${c_V}^0$ corresponds to the specific heat of a free electron gas, which is (in rydbergs)

$${c_V}^0 = \frac{4\hbar^2 e^{-4} k^2 T}{9\pi m}\, {r_s}^2. \tag{9.13}$$

Problems

1. Solving Bloch's equation, evaluate the density matrix of a free particle in a one-dimensional box of length L.
2. In the above problem, show that, for large L, the distribution function assumes the form

$$f(x) = \frac{1}{\lambda}\left[1 - \exp\left(\frac{-4\pi x^2}{\lambda^2}\right)\right].$$

3. The eigenfunctions and the eigenvalues of a rigid rotor are given by

$$\psi_{lm}(\theta, \varphi) = \left[\frac{2l+1}{4\pi}\frac{(l-m)!}{(l+m)!}\right]^{1/2} P_l^m(\cos\theta)e^{im\varphi};$$
$$\varepsilon_l = \frac{\hbar^2}{2A} l(l+1), \qquad l = 0, 1, \ldots; \quad m = -l, \ldots, l.$$

Evaluate the density matrix and show that the partition function is given by

$$Z_r = \sum (2l+1)\exp\left[-\left(\frac{\hbar^2}{2A}\right)\beta l(l+1)\right],$$

where A is the transverse moment of inertia.

4. Using the Euler–Maclaurin summation formula, derive the approximate expression

$$Z_r = \frac{2AkT}{h^2} + \frac{1}{3} + \frac{1}{15}\frac{h^2}{2AkT} + \cdots.$$

5.* Express the partition function in the form

$$Z = \int_{-\infty}^{\infty} dE\, e^{-\beta E} N(E),$$

where the density of states $N(E)$ is

$$N(E) = \sum_{E'} \langle E'|\delta(H - E)|E'\rangle$$
$$= \frac{1}{\pi}\lim_{\varepsilon\to 0+} \mathrm{tr}\left(\frac{1}{H - E - i\varepsilon}\right).$$

6. Show that

$$(e^{-\beta H})_{kk} \geq e^{-\beta H_{kk}}.$$

7. Examine the property of the limit of $K_0(\mathbf{r}\beta;\mathbf{r}'\beta')$ when $\beta' \to \beta-$.

8.* For a density matrix $\rho = e^{-\beta H}$ with $H = H_0 + \Phi$, define a Laplace transform:

$$R(E) = -\int_0^\infty e^{\beta E}\rho(\beta)\,d\beta$$

in the region where the real part of E is less than the smallest eigenvalue of H. Using a boundary condition $\rho(0) = 1$, show that the Møller operator

$$G(E) = R(E)(E - K)$$

satisfies

$$G(E) = 1 + (E - H_0)^{-1}\Phi G(E).$$

9.* Define a two-body scattering operator t_α by

$$t_\alpha = \phi_\alpha + \phi_\alpha(E - H_0)^{-1}t_\alpha,$$

where ϕ_α represents the potential of a pair of particles. Show that (*2*)

$$G(E) = 1 + (E - H_0)^{-1}\sum_\alpha t_\alpha G_\alpha,$$

$$G_\alpha = 1 + (E - H_0)^{-1}\sum_\beta t_\beta G_\beta.$$

10.* By perturbation, derive (*18*)

$$G(E) = 1 + (E - H_0)^{-1}\sum t_\alpha + (E - H_0)^{-1}\sum_\alpha t_\alpha(E - H_0)^{-1}\sum_{\beta\neq\alpha} t_\beta + \cdots.$$

11.* By using the approximation $G = 1 + (E - H_0)^{-1}t$, express the second virial coefficient in the following form:

$$B = -\frac{v}{(\pi mkT)^{3/2}}\frac{1}{2\pi i}\int d\mathbf{p}\int dE\, e^{-\beta E}\frac{\langle p|t|p\rangle}{(E - H_0)^2}.$$

12.* Determine the propagator of a free electron in a homogeneous magnetic field (*19*).

GENERAL REFERENCES

E. W. Montroll, *in* "The Theory of Neutral and Ionized Gases" (C. De Witt and J. F. Detoeff, eds.). Wiley, New York, 1960.

R. Balian, C. Bloch, and C. De Dominicis, *Nuclear Phys.* **25**, 529 (1961); **27**, 294 (1961).

A. Isihara, *in* "Lectures in Theoretical Physics," Vol. XA (A. D. Barut and W. E. Brittin, eds.). Gordon and Breach, New York, 1968.

REFERENCES

1. K. Husimi, *Proc. Phys. Math. Soc. Japan, Ser. 3* **22**, 264 (1940).
2. M. Goldberger and E. N. Adams, *J. Chem. Phys.* **20**, 240 (1952).
3. A. J. F. Siegert, *J. Chem. Phys.* **20**, 572 (1952).
4. T. Matsubara, *Progr. Theoret. Phys.* **14**, 351 (1955).
5. K. A. Brueckner, *Phys. Rev.* **100**, 36 (1955).
6. K. M. Watson, *Phys. Rev.* **103**, 489 (1956).
7. J. Goldstone, *Proc. Roy. Soc. (London) Ser. A* **239**, 267 (1957).
8. N. N. Hugenholtz, *Physica* **23**, 481 (1957).
9. J. Hubbard, *Proc Roy. Soc. (London) Ser. A* **243**, 336 (1957).
10. C. Bloch and C. De Dominicis, *Nuclear Phys.* **9**, 459 (1958).
11. J. E. Moyal, *Proc. Camb. Philos. Soc.* **45**, 99 (1949).
12. M. Gell-Mann and K. A. Brueckner, *Phys. Rev.* **106**, 364 (1957).
13. E. W. Montroll and J. C. Ward, *Phys. Fluids* **1**, 55 (1958).
14. A. Isihara, *Progr. Theoret. Phys. Suppl.* **44**, 1 (1969).
15. A. Isihara, *Phys. Rev.* **172**, 166 (1968).
16. M. Gell-Mann, *Phys. Rev.* **106**, 369 (1957).
17. D. Pines and P. Nozière, "The Theory of Quantum Liquids." Benjamin, New York, 1966.
18. K. M. Watson, *Phys. Rev.* **103**, 489 (1956).
19. A. Isihara and M. Wadati, *Phys. Rev.* **A1**, 318 (1970).

11

Propagator Methods for Distribution Functions

We shall discuss a propagator formalism for the evaluation of the reduced distribution functions, especially the pair distribution function, of a quantum mechanical many-body system in equilibrium. The reduced distribution functions are obtained from the diagonal elements of the density matrix of the entire system and represent the probability distributions of particles. For instance, the pair distribution function gives the probability of finding two particles at given coordinates. As in the classical cases, it is a function of distance when the system is homogeneous, isotropic, and in equilibrium.

From the distribution functions, one can obtain information concerning the spatial correlations of particles. In addition, one can evaluate the energy, the equation of state, and other average quantities.

We shall first discuss the linked cluster expansion of the distribution functions. We then introduce a formula for the pair distribution function which is obtained by resumming the expansion corresponding to chain diagrams. As far as chain diagrams are concerned, the formula is quite general. However, we shall be concerned mainly with collective coupling effects on spatial correlations, since the chain diagram approximation is very effective in this respect and the results are interesting. Improvements upon this approximation are possible, but are beyond the scope of this book.

11.1. Linked Cluster Expansions of Distribution Functions

The reduced distribution function of l particles in the canonical ensemble may be defined in terms of the propagator as follows:

$$\rho_l^{(N)}(\mathbf{r}^l) = \frac{N!}{(N-l)!\,Z_N} \int \cdots \int K(\mathbf{r}^N\beta;\, \mathbf{r}^N 0)\, d\mathbf{r}_{l+1} \cdots d\mathbf{r}_N, \tag{1.1}$$

where the superscript (N) is introduced to indicate a system with a fixed number N of particles. In the grand ensemble, the distribution function is (l)

$$\rho_l(\mathbf{r}^l) = \frac{1}{\Xi} \sum_{N=l}^{\infty} \rho_l^{(N)}(\mathbf{r}^l) Z_N z^N. \tag{1.2}$$

The distribution function is normalized such that its integration over $\mathbf{r}_1, \ldots, \mathbf{r}_l$ yields the following grand ensemble average:

$$\begin{aligned} \int \rho_l(\mathbf{r}^l)\, d\mathbf{r}^l &= \frac{1}{\Xi} \sum_N N(N-1) \cdots (N-l+1) Z_N z^N \\ &= \langle N(N-1) \cdots (N-l+1) \rangle. \end{aligned} \tag{1.3}$$

We shall derive a linked cluster expansion for the pair distribution function (p.d.f.) $\rho_2(\mathbf{r}_1, \mathbf{r}_2)$. For this purpose, let us consider $K(\mathbf{r}^N\beta;\, \mathbf{r}^N 0)$. When expanded as in Eq. (4.15) of Chapter 10, it yields two different types of terms: First, terms in which $\mathbf{r}_1$ and $\mathbf{r}_2$ appear in separate diagrams, and second, those in which $\mathbf{r}_1$ and $\mathbf{r}_2$ are in the same connected diagram. There are terms in these groups that are independent of $\mathbf{r}_1$ and $\mathbf{r}_2$ but depend on the other particles. These terms, when integrated, produce products of b_l's. The contributions to the p.d.f. of the first and the second types may be obtained as follows (l). The b_l's are defined by Eq. (4.22) of Chapter 10.

All the terms in the expansion of $K(\mathbf{r}^N\beta;\, \mathbf{r}^N 0)$ depend only on the relative distances. Therefore, upon integrating over $\mathbf{r}_3, \ldots, \mathbf{r}_N$, the terms that contain $\mathbf{r}_1$ and not $\mathbf{r}_2$ will be independent of $\mathbf{r}_1$. The result of integration should be V^{-1} times that which would have resulted from an integration over $\mathbf{r}_1$. Thus, if $\mathbf{r}_1$ is included in an $(l+1)$-toron we expect $V^{-1} b_{l+1}$. However, we must multiply this by $(l+1)$ because the particle 1 may be put in any of the $l+1$ loops of the $(l+1)$-toron. Various $(l+1)$-toron diagrams involving $\mathbf{r}_1$ in different locations yield, upon integration over $\mathbf{r}_1$, the same b_{l+1}.

Similarly, an $(m+1)$-toron which involves $\mathbf{r}_2$ gives $(m+1) b_{m+1} V^{-1}$. Thus, the contribution to the p.d.f. from the diagrams containing either $\mathbf{r}_1$ or $\mathbf{r}_2$ is

$$Z_N^{-1} \sum_{l=0}^{N-2} \sum_{m=0}^{N-2-l} (l+1) b_{l+1} (m+1) b_{m+1} V^{-2} \sum_{\{s_l\}} \sum_N \prod_{l=m-2} b_l^{s_l} s_l!. \tag{1.4}$$

We may express the last factor by a Cauchy integral:

$$\sum \prod \frac{b_t^{s_t}}{s_t!} = \frac{1}{(2\pi i)} \oint \frac{\exp(\sum w^t b_t)\, dw}{w^{N-l-m-1}}. \tag{1.5}$$

Multiplying $z^N Z_N$ by Eq. (1.4) and summing over all N, we arrive at

$$\begin{aligned} &\frac{V^{-2}}{2\pi i} \sum_{l,m} \oint (l+1) b_{l+1} w^{l+1} (m+1) b_{m+1} w^{m+1} \exp(\textstyle\sum w^t b_t) \frac{dw}{w-z} \\ &= \left[\frac{1}{V} \sum (l+1) b_{l+1} z^{l+1}\right]^2 \exp\left(\sum_t z^t b_t\right) \\ &= n^2 \Xi, \end{aligned} \tag{1.6}$$

where n is the number density and where we have used for the present b_l

$$\sum_{l=0}^{\infty} (l+1) b_{l+1} z^{l+1} = \sum_{l=0}^{\infty} l b_l z^l = nV,$$

and also

$$\Xi = \exp(\textstyle\sum_t z^t b_t)$$

[see Eq. (6.6) of Chapter 10]. Thus, the total contribution from the graphs which contain separately either $\mathbf{r}_1$ or $\mathbf{r}_2$ is simply n^2.

Let us now consider $(l+2)$-particle diagrams which contain both $\mathbf{r}_1$ and $\mathbf{r}_2$. Integration of $K(\mathbf{r}^N\beta; \mathbf{r}^N 0)$ over the coordinates of l particles except $\mathbf{r}_1$ and $\mathbf{r}_2$ yields

$$b_{l+2}(\mathbf{r}_1, \mathbf{r}_2) \sum \prod \frac{b_t^{s_t}}{s_t!}. \tag{1.7}$$

where $b_{l+2}(\mathbf{r}_1, \mathbf{r}_2)$ is a function of $\mathbf{r}_1, \mathbf{r}_2$, and $l+2$, satisfying the condition

$$\int b_{l+2}(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_1\, d\mathbf{r}_2 = (l+2)(l+1) b_{l+2}. \tag{1.8}$$

This results because there are $(l+2)(l+1)$ ways of putting $\mathbf{r}_1$ and $\mathbf{r}_2$ in an $(l+2)$-connected diagram.

Multiplying Eq. (1.7) by z^{N+l+2}/Ξ and summing over all N, we obtain from the diagrams containing an exchange of $\mathbf{r}_1$ and $\mathbf{r}_2$ the following contribution to the p.d.f.:

$$\sum_{l=2} b_l(\mathbf{r}_1, \mathbf{r}_2) z^l.$$

Thus, adding this to Eq. (1.6), we arrive at

$$\rho_2(\mathbf{r}_1, \mathbf{r}_2) = n^2 + \sum_{l=2}^{\infty} b_l(\mathbf{r}_1, \mathbf{r}_2) z^l, \tag{1.9}$$

that is, a cluster expansion for the p.d.f.

We can check the normalization of both sides of this equation in the following way. Upon integration, we expect that

$$\begin{aligned} \int \rho_2(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_1\, d\mathbf{r}_2 &= \langle N(N-1)\rangle \\ &= \langle N\rangle^2 + \sum_l l(l-1) b_l z^l. \end{aligned} \tag{1.10}$$

On the other hand, from Eq. (6.6) of Chapter 10, we have

$$\langle N\rangle = \sum_l l b_l z^l,$$

and thus

$$\frac{\partial(\langle N\rangle)}{\partial(\log z)} = \sum_l l^2 b_l z^l,$$

or

$$\sum_{l=2}^{\infty} l^2 b_l z^l = \langle N^2\rangle - \langle N\rangle^2.$$

Thus

$$\begin{aligned} \frac{\partial(\langle N\rangle)}{\partial(\log z\rangle)} - \langle N\rangle &= \langle N(N-1)\rangle - \langle N\rangle^2 \\ &= \sum_{l=2}^{\infty} l(l-1) b_l z^l. \end{aligned} \tag{1.11}$$

This confirms that Eq. (1.10) is correct.

11.2. Distribution Functions of Ideal Quantum Gases

Let us now apply Eq. (1.9) to an ideal gas. For this case, $b_l(\mathbf{r}_1, \mathbf{r}_2)$ is obtained from an l-toron involving $\mathbf{r}_1$ and $\mathbf{r}_2$. Let us consider a toron which has s loops in the path from $\mathbf{r}_1$ to $\mathbf{r}_2$ and t loops in the path from $\mathbf{r}_2$ to $\mathbf{r}_1$ such that $t + s = l$.

The propagator that carries a single particle from $(\mathbf{r}', \beta')$ to $(\mathbf{r}, \beta)$ is

$$\begin{aligned} K_0(\mathbf{r}\beta; \mathbf{r}'\beta') &= \left[\left(\frac{2\hbar^2\pi}{m}\right)(\beta - \beta')\right]^{-3/2} \exp\left[-\frac{m(\mathbf{r}-\mathbf{r}')^2}{2\hbar^2(\beta-\beta')}\right] \\ &= \frac{1}{h^3} \int \exp\left[-\frac{p^2(\beta-\beta')}{2m} + i\hbar^{-1}\mathbf{p}\cdot(\mathbf{r}-\mathbf{r}')\right] d\mathbf{p}. \end{aligned} \tag{2.1}$$

The propagator that brings $\mathbf{r}_1$ to $\mathbf{r}_2$ after s cycles is obtained by replacing

respectively $\mathbf{r}$, β in this expression by $\mathbf{r}_2$, $s\beta$ and also $\mathbf{r}'$, β' by $\mathbf{r}_1$, 0. A similar expression is obtained for the propagator that brings the particle from $\mathbf{r}_2$ to $\mathbf{r}_1$ after t cycles. Thus,

$$b_l(\mathbf{r}_1, \mathbf{r}_2) = \sum_{s+t=l} \frac{(\pm)^{s+t+1}}{h^6} \int \exp\left[-\frac{s\beta p^2}{2m} + i\hbar^{-1}\mathbf{p}\cdot(\mathbf{r}_2 - \mathbf{r}_1)\right] d\mathbf{p}$$
$$\times \int \exp\left[-\frac{t\beta q^2}{2m} + i\hbar^{-1}\mathbf{q}\cdot(\mathbf{r}_1 - \mathbf{r}_2)\right] d\mathbf{q}, \tag{2.2}$$

where $l = s + t$, and the sign factor is plus for bosons and minus for fermions.

Introducing Eq. (2.2) into Eq. (1.9) and summing the series in z, we obtain

$$\rho_2{}^0(\mathbf{r}_1 - \mathbf{r}_2) = n^2 \pm \left|\frac{1}{h^3}\int \frac{z\exp(-\beta p^2/2m)\exp[i\hbar^{-1}\mathbf{p}\cdot(\mathbf{r}_2 - \mathbf{r}_1)]}{1 \mp z\exp(-\beta p^2/2m)}\, d\mathbf{p}\right|^2, \tag{2.3}$$

where the upper (lower) sign corresponds to bosons (fermions) and where z is determined from

$$n^{(0)} = \frac{1}{h^3}\int \frac{z\exp(-\beta p^2/2m)\, d\mathbf{p}}{1 \mp z\exp(-\beta p^2/2m)}. \tag{2.4}$$

The p.d.f. obtained from Eq. (2.3) is illustrated in Fig. 11.1 as a function of

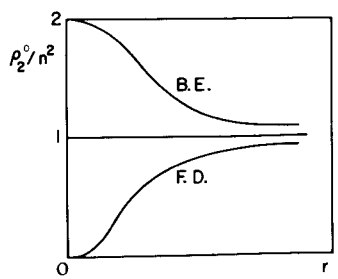


Fig. 11.1. The pair distribution function for ideal quantum gases.

$|\mathbf{r}_2 - \mathbf{r}_1| = r$. It is interesting to see the different behaviors of the p.d.f.'s for ideal fermions and ideal bosons near the origin.

The p.d.f. of an ideal Bose gas is determined by the function

$$G_{3/2}(z, r) = \frac{\lambda^3}{h^3}\int \frac{\exp(i\hbar^{-1}\mathbf{p}\cdot\mathbf{r})}{z^{-1}[\exp(\beta p^2/2m)] - 1}\, d\mathbf{p}$$
$$= \sum_l \frac{z^l}{l^{3/2}} \exp\left[-\frac{\pi r^2}{l\lambda^2}\right], \tag{2.5}$$

where the integral is real. This function generalizes $G_s(\lambda)$ of Eq. (3.16)

of Chapter 4. Near the condensation temperature of an ideal Bose gas, it is approximately given by

$$G_{3/2}(z, r) = \frac{\lambda}{r}\exp\left[-2\pi^{1/2}\alpha^{1/2}\frac{r}{\lambda}\right], \tag{2.6}$$

where λ is the thermal de Broglie wavelength $(h^2/2\pi mkT)^{1/2}$ and $\alpha = -\log z$.

Using the expansions given by Eq. (4.12) of Chapter 4, we arrive at the expression

$$\frac{\rho_2^0(r)}{n^2} - 1 = (-3.54\alpha^{1/2} + 2.61 + 1.46\alpha)^{-2}\frac{\lambda^2}{r^2}\exp\left[-\frac{4\pi^{1/2}\alpha^{1/2}r}{\lambda}\right], \tag{2.7}$$

which is valid near the condensation temperature T_0. Here,

$$\alpha^{1/2} = \frac{1.105(T - T_0)}{T_0}. \tag{2.8}$$

As we approach the condensation temperature, the spatial correlation of the Bose particles becomes more and more long-ranged.

The pair distribution function of an ideal Fermi gas is determined by the function

$$G_{3/2}(z, r) = \frac{\lambda^3}{h^3}\int \frac{\exp(i\hbar^{-1}\mathbf{p}\cdot\mathbf{r})}{z^{-1}[\exp(\beta p^2/2m)] + 1}\, d\mathbf{p}$$
$$= -\sum_l \frac{(-z)^l}{l^{3/2}}\exp\left[-\frac{\pi r^2}{l\lambda^2}\right]. \tag{2.9}$$

This function again generalizes the corresponding Fermi function that appeared in Eq. (3.16) of Chapter 4. Although we are using the same notation $G_{3/2}$ for both Bose and Fermi cases, the difference is clear.

The integral may be evaluated for low temperatures as follows. We choose new variables $x = p_0(p - p_0)/mkT$ and $u = (e^x + 1)^{-1}$ and write $v = mkT/\hbar p_0$. We find, for $T \to 0$,

$$G_{3/2}(z, r) = -\frac{\lambda^3}{h^3}\frac{h^2}{\pi r}\frac{\partial}{\partial r}\int_0^\infty \frac{\cos(pr/\hbar)}{\exp\beta[(p^2/2m) - (p_0{}^2/2m)] + 1}\, dp$$
$$= -\frac{4\pi\lambda^3}{h^3}\frac{h^3}{r}\frac{\partial}{\partial r}\left[\frac{1}{r}\left(\sin\frac{p_0 r}{\hbar}\right)\int_0^\infty e^{ivxr}\, du\right]$$
$$= -\frac{\lambda^3}{2\pi^2 r}\frac{\partial}{\partial r}\left|\left(\sin\frac{p_0 r}{\hbar}\right)\left(\frac{\pi v}{\sinh \pi v r}\right)\right|. \tag{2.10}$$

In particular, at $T = 0$, Eq. (2.10) reduces to

$$G_{3/2}(z, r) = \frac{\lambda^3}{2\pi^2} \frac{\sin(k_0 r) - k_0 r \cos(k_0 r)}{r^3}, \tag{2.11}$$

where $2\pi k_0$ is the wave number at the Fermi surface.

11.3. Triplet Distribution Function

By a similar consideration of diagrams, one can obtain the triplet and higher-order distribution functions. For example, the diagrams related to the triplet distribution function may be classified into the following categories:

(a) The three representative particles are on three independent graphs. Clearly, the total contribution of this type of graph is

$$I_1 = \rho_1(\mathbf{r}_1)\rho_1(\mathbf{r}_2)\rho_1(\mathbf{r}_3). \tag{3.1}$$

(b) Two of the three particles are on the same toron, while the third is on an independent toron. We have the contribution

$$\pm[\rho_1(\mathbf{r}_1)I_2(\mathbf{r}_2, \mathbf{r}_3) + \rho_1(\mathbf{r}_2)I_2(\mathbf{r}_3, \mathbf{r}_1) + \rho_1(\mathbf{r}_3)I_2(\mathbf{r}_1, \mathbf{r}_2)],$$

where $\rho_1(\mathbf{r}_i)$ is the singlet distribution function and

$$\begin{aligned} I_2(\mathbf{r}_1, \mathbf{r}_2) &= \left\{\frac{1}{h^3}\int \frac{\exp[ih^{-1}\mathbf{p}\cdot(\mathbf{r}_2 - \mathbf{r}_1)]}{z^{-1}[\exp(\beta p^2/2m)] \mp 1}\, d\mathbf{p}\right\}^2. \\ &= [\lambda^{-3}G_{3/2}(z, |\mathbf{r}_2 - \mathbf{r}_1|)]^2. \end{aligned} \tag{3.2}$$

(c) The three particles are on the same toron. The contribution from an N-toron for fermions is represented by the integral

$$\begin{aligned} I_3^{(N)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = &\sum_{l+s+t=N} (-)^{l+s+t+1} h^{-9} \\ &\times \int \exp\left[-\frac{l\beta p_1^2}{2m} + ih^{-1}\mathbf{p}_1\cdot(\mathbf{r}_2 - \mathbf{r}_1)\right] d\mathbf{p}_1 \\ &\times \int \exp\left[-\frac{s\beta p_2^2}{2m} + ih^{-1}\mathbf{p}_2\cdot(\mathbf{r}_3 - \mathbf{r}_2)\right] d\mathbf{p}_2 \\ &\times \int \exp\left[-\frac{t\beta p_3^2}{2m} + ih^{-1}\mathbf{p}_3\cdot(\mathbf{r}_1 - \mathbf{r}_3)\right] d\mathbf{p}_3. \end{aligned} \tag{3.3}$$

Multiplying this by z^N and the weight factor 2, which comes from the permutation of the three positions in a ring, and summing over all N, we obtain the contribution:

$$\begin{aligned} I_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) &= \sum_N 2I_3^{(N)}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)z^N \\ &= 2\prod_{j\neq i}^{3}\left\{\frac{1}{h^3}\int \frac{z\exp(-\beta p_i^2)\exp[ih^{-1}\mathbf{p}_i\cdot(\mathbf{r}_j - \mathbf{r}_i)]}{1 \mp z\exp(-\beta p_i^2/2m)}\, d\mathbf{p}_i\right\}. \end{aligned} \tag{3.4}$$

We arrive at

$$\rho_3{}^0(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = n^3 \pm n[I_2(\mathbf{r}_1, \mathbf{r}_2) + I_2(\mathbf{r}_2, \mathbf{r}_3) + I_2(\mathbf{r}_3, \mathbf{r}_1)] + I_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3). \tag{3.5}$$

The triplet distribution function satisfies the normalization condition

$$\int \rho_3{}^0(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)\, d\mathbf{r}^3 = \langle N(N-1)(N-2)\rangle. \tag{3.6}$$

This can be confirmed easily by noting

$$\begin{aligned} \int I_3\, d\mathbf{r}_1\, d\mathbf{r}_2\, d\mathbf{r}_3 &= \frac{\partial^2(\langle N\rangle)}{\partial[(\log z)^2]} - \langle N\rangle \mp 3\int I_2\, d\mathbf{r}_1\, d\mathbf{r}_2 \\ &= \langle N^3\rangle - 3\langle N^2\rangle\langle N\rangle + 2\langle N\rangle^3 - \langle N\rangle \\ &\quad - 3[\langle N^2\rangle - \langle N\rangle^2 - \langle N\rangle], \end{aligned} \tag{3.7}$$

and

$$\pm\int I_2(\mathbf{r}_1, \mathbf{r}_2)\, d\mathbf{r}_1\, d\mathbf{r}_2 = \langle N^2\rangle - \langle N\rangle^2 - \langle N\rangle.$$

11.4. Chain Diagram Approximation

Let us calculate $b_l(\mathbf{r}_1, \mathbf{r}_2)$ in the lowest order in the interaction parameter e^2. It is convenient to consider a process with the following steps corresponding to Eq. (7.10) of Chapter 10 (see Fig. 11.2): (1) a particle propagates from

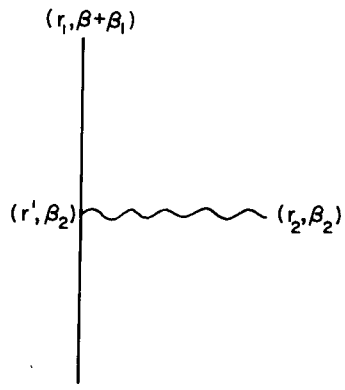


Fig. 11.2. The propagation for G_1.

$(\mathbf{r}_1, \beta_1)$ to $(\mathbf{r}', \beta_2)$, (2) interacts at $\mathbf{r}'$ with another particle which is at a point $\mathbf{r}_2$, and then (3) propagates from $(\mathbf{r}', \beta_2)$ back to the point $(\mathbf{r}_1, \beta + \beta_1) = (\mathbf{r}_1, \beta_1)$, completing a cycle of β.

The above three steps are expressed by a product of corresponding propagators and the interaction as follows:

$$K_0(\mathbf{r}_1, \beta + \beta_1; \mathbf{r}', \beta_2)\phi(|\mathbf{r}_2 - \mathbf{r}'|)K_0(\mathbf{r}'\beta_2; \mathbf{r}_1\beta_1). \tag{4.1}$$

Integrating this over all intermediate coordinates $\mathbf{r}'$, we get the result of Eq. (7.11) of Chapter 10:

$$F_1(\mathbf{r}_2 - \mathbf{r}_1; \beta_2 - \beta_1) = \frac{1}{h^3}\int G_1(\mathbf{q}; \beta_2 - \beta_1)u(q) \times \exp[i\hbar^{-1}\mathbf{q}\cdot(\mathbf{r}_2 - \mathbf{r}_1)]\, d\mathbf{q}. \tag{4.2}$$

where G_1 is defined by Eq. (7.10) of Chapter 10.

As before, F_1 may be expanded in a Fourier series in $\beta_2 - \beta_1$ with a period β. We find

$$F_1(\mathbf{r}_2 - \mathbf{r}_1; \beta_2 - \beta_1) = -\frac{1}{\beta h^3}\sum_j \int d\mathbf{q}\, u(q)\lambda_j(q) \times \exp[i\hbar^{-1}\mathbf{q}\cdot(\mathbf{r}_2 - \mathbf{r}_1)] \exp\left[\frac{2\pi i j(\beta_2 - \beta_1)}{\beta}\right], \tag{4.3}$$

where

$$\lambda_j(q) = z\lambda^{-3}\int \exp\left[-\frac{q^2\alpha(\beta - \alpha)}{2m\beta}\right] \exp\left[\frac{2\pi i j\alpha}{\beta}\right] d\alpha \tag{4.4}$$

$$= z\lambda^{-3}\beta\left[1 + \frac{q^2\beta}{2\pi^2 j^2(2m)} - \frac{12q^4\beta^2}{(2\pi j)^4(2m)^2} + \cdots\right], \qquad j \neq 0 \tag{4.5}$$

λ being the thermal de Broglie wavelength.

The process which led to Eq. (4.3) may be repeated twice. The corresponding propagator is

$$F_2(\mathbf{r}_3\beta_3; \mathbf{r}_1\beta_1) = \frac{1}{\beta h^3}\sum_j \int d\mathbf{q}\, u^2(q)\lambda_j^2(q) \times \exp[i\hbar^{-1}\mathbf{q}\cdot(\mathbf{r}_3 - \mathbf{r}_1)] \exp\left[\frac{2\pi i j(\beta_3 - \beta_1)}{\beta}\right]. \tag{4.6}$$

Repetition of the same process gives us the propagator for l-member chains:

$$F_l(\mathbf{r}_2\beta_2; \mathbf{r}_1\beta_1) = \frac{1}{\beta h^3}\sum_j \int d\mathbf{q}\, [-u(q)]^{l-1}\lambda_j^l(q) \times \exp[i\hbar^{-1}\mathbf{q}\cdot(\mathbf{r}_2 - \mathbf{r}_1)] \exp\left[\frac{2\pi i j(\beta_2 - \beta_1)}{\beta}\right], \tag{4.7}$$

where l represents the total number of l-torons.

Summing the product in Eq. (4.7) over all l, we obtain

$$\rho_2(r) - n^2 = \mp I_2(r) - \frac{1}{\beta h^3}\sum_j \int \frac{u(q)\lambda_j^2}{1 + \lambda_j u(q)} \exp(i\hbar^{-1}\mathbf{q}\cdot\mathbf{r})\, d\mathbf{q}. \tag{4.8}$$

The pair distribution function given by this formula is based on chain diagrams. The chain diagram contribution is important in the sense that collective couplings of the particles can be seen in the result.

Equation (4.8) is applicable to a simple chain of toron graphs if the λ_j represent the eigenvalues of $G(q, \alpha)$ of Eq. (8.1) of Chapter 10 instead of G_1. If exchange graphs due to quantum statistics are taken into consideration, or if the distances are short, the diagrams in which 1 and 2 are exchanged must be included in our consideration (Fig. 11.3). Indeed, due to these exchange

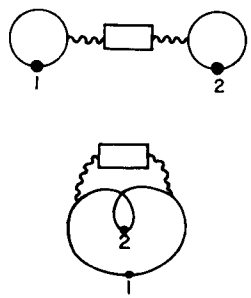


Fig. 11.3. Classical (top) and quantum (bottom) chain diagrams.

diagrams, one can show that the p.d.f. for fermions vanishes at $r = 0$ if the spins are the same. The result based on chain diagrams is given by (*2*)

$$\rho_2(r) - n^2 = \mp I_2(r) - \frac{1}{h^3\beta}\sum_j \int \frac{u(q)}{1 + \lambda_j u(q)} \times [\lambda_j^2 \mp \lambda_j^2(q, r)] \exp(i\hbar^{-1}\mathbf{q}\cdot\mathbf{r})d\mathbf{q}, \tag{4.9}$$

where

$$\lambda_j(q, r) = \frac{1}{h^3}\iint f^{(0)}(p)[1 \mp f^{(0)}(\mathbf{p} + \mathbf{q})] \times \exp\frac{\alpha[p^2 - (\mathbf{p} + \mathbf{q})^2]}{2m} \exp\left(i\hbar^{-1}\,\mathbf{p}\cdot\mathbf{r} + \frac{2\pi i j\alpha}{\beta}\right) d\alpha\, d\mathbf{p}. \tag{4.10}$$

On the right-hand sides of Eqs. (4.9) and (4.10), the upper and lower signs correspond respectively to fermions and bosons. For fermions, we use the Fermi distribution $f^{(0)}$ and the eigenvalues given by Eq. (8.5) of Chapter 10 [cf. Eq. (6.5) of this chapter]. $I_2(r)$ has been defined by Eq. (3.2). Equation (4.9) is obtained by considering the following processes: (*i*) A propagation

from $(\mathbf{r}_1, 0)$ with momentum $\mathbf{p}_1 + \mathbf{q}$, followed by a propagation with momentum $\mathbf{p}_1$ after an emission of momentum of $\mathbf{q}$. (*ii*) A propagation from $(\mathbf{r}_2, 0)$ with momentum $\mathbf{p}_2$, followed by a propagation with momentum $\mathbf{p}_2 + \mathbf{q}$ after an absorption of momentum $\mathbf{q}$.

If the first process corresponds to $(\mathbf{r}_1, 0) \rightarrow (\mathbf{r}_1, \beta)$ and the second one to $(\mathbf{r}_2, 0) \rightarrow (\mathbf{r}_2, \beta)$, the two combined yield the first term in the square brackets of Eq. (4.9). If the two processes are for $(\mathbf{r}_1, 0) \rightarrow (\mathbf{r}_2, \beta)$ and $(\mathbf{r}_2, 0) \rightarrow (\mathbf{r}_1, \beta)$, we introduce a transformation of the momentum variable of the particle 2 from $\mathbf{p}_2 + \mathbf{q}$ to $-\mathbf{p}_2$. Due to this transformation, the phase factor becomes $\exp[i(\mathbf{p}_1 \cdot \mathbf{r} + \mathbf{p}_2 \cdot \mathbf{r} + \mathbf{q} \cdot \mathbf{r})/\hbar]$. We then obtain the second term in the brackets of Eq. (4.9).

Equation (4.9) is the result based on diagrams with a single effective interaction. One can generalize the formula to the cases with more interactions. For instance, the p.d.f. in the ring diagram approximation has been obtained (*2*).

11.5. Classical Electron Gas

By using the chain diagram approximation, one can show that the p.d.f. of a classical electron gas is determined by the well-known Debye screening potential. For this purpose, all we need is to introduce the eigenvalues for Boltzmann particles given by Eq. (7.18) of Chapter 10 into Eq. (4.8). We note first that, at high temperatures, the following approximation holds:

$$\lambda_j = z\lambda^{-3}\beta\,\delta_{j,0} = n\beta\delta_{j,0}, \tag{5.1}$$

where n is the number density and δ is a Kronecker delta function. The Fourier transform of the Coulomb potential has been given by Eq. (7.2) of Chapter 10:

$$u(q) = \hbar^2 \frac{4\pi e^2}{q^2}. \tag{5.2}$$

Introducing Eqs. (5.1) and (5.2) into Eq. (4.8), we find

$$\rho_2(r) - n^2 = -\frac{z\kappa^2}{\lambda^3\hbar^3}\int \frac{\exp[i\hbar^{-1}\mathbf{q}\cdot(\mathbf{r}_2 - \mathbf{r}_1)]}{q^2 + \kappa^2}\, d\mathbf{q}, \tag{5.3}$$

where κ^{-1} is the Debye length, given by

$$\kappa = \left(\frac{4\pi n e^2}{kT}\right)^{1/2}. \tag{5.4}$$

Performing the $\mathbf{q}$-integration of Eq. (5.3), we obtain

$$\rho_2(r) = n^2 - \frac{n\kappa^2 e^{-\kappa r}}{4\pi r}. \tag{5.5}$$

We conclude from this result that the total contribution of the classical chain diagrams with 1 and 2 at the ends may be represented by an effective interaction $e^2e^{-\kappa r}/r$. It is instructive to rewrite Eq. (5.5) in the following reduced form:

$$\frac{\rho_2(r)}{n^2} - 1 = g(r) = -\varepsilon\frac{e^{-x}}{x}, \tag{5.6}$$

where $g(r)$ is the correlation function, the parameter ε is called the plasma parameter, and x is a reduced distance variable; ε and x are defined by

$$\varepsilon = \beta e^2\kappa, \qquad x = \kappa r. \tag{5.7}$$

Note that in the reduced form (5.6) the p.d.f. has been evaluated to the first order in ε. Therefore, Eq. (5.6) may be called the first-order Debye–Hückel distribution function (*2*).

The parameter ε is small for a classical electron gas. For instance, ε is of the order 10^{-2} for $n \sim 10^{13}$ cm^{-3} and $T \sim 10^4$ °K.

The right-hand side of Eq. (5.6) diverges at $x = 0$. In fact, Eq. (5.6) is valid only at large distances. It is obvious from Eq. (4.9) that, at short distances, exchange diagrams play important roles. That is, the divergence can be eliminated by quantum effects. Even in the classical approximation, one can remove the short-distance divergence by taking into account infinite effective chains between 1 and 2. Remembering that a single chain corresponds to a screening potential and that l effective lines have a degeneracy weight factor $1/l!$, we arrive at

$$\rho_2(r) = n^2 e^{-\Psi(r)/kT}, \qquad \Psi(r) = \frac{e^2e^{-\kappa r}}{r}. \tag{5.8}$$

$\Psi(r)$ is the Debye–Hückel potential. The $\rho_2(r)$ of Eq. (5.8) is called the "watermelon result" because the graphs with many chains of interactions resemble watermelons. It is convergent at $r = 0$ (Fig. 11.4).

Before examining Eq. (5.8), we illustrate the Debye-Hückel distribution function (5.8) for several different values of the plasma parameter ε. As can be seen, the curves approach the asymptotic value 1 rapidly as the plasma

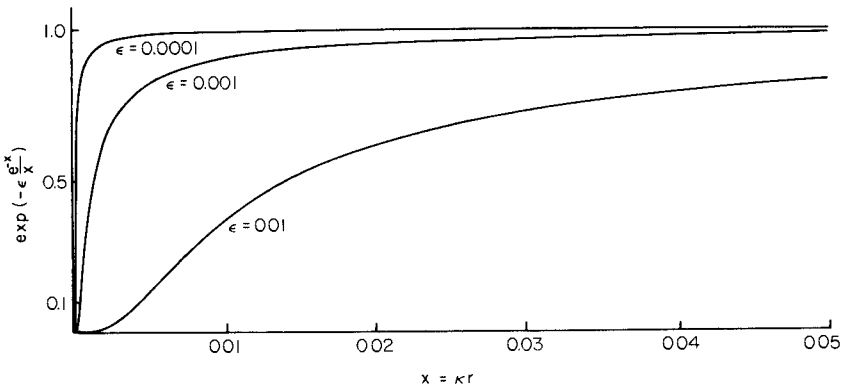


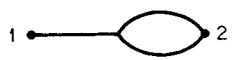
Fig. 11.4. The pair distribution function obtained by superposing chain diagrams (watermelon approximation). The abscissa is a reduced distance κr.

parameter is reduced. When the distance r is of the order of κ^{-1}, the particle correlations disappear. Thus, for practical purposes, ε is a good quantity for expressing the effectiveness of particle correlation in a classical plasma.

Equation (5.8) contains corrections to the first-order Debye-Hückel distribution function. When higher-order terms in ε appear in $\rho_2(r)$, it becomes necessary to examine the results very carefully. As a matter of fact, Eq. (5.8) is not satisfactory to the order ε^2, because some other diagrams contribute to the same order. In addition, when higher-order terms in ε are involved, it generally becomes necessary to take quantum effects into consideration because $\lambda\kappa$, the de Broglie wavelength of the electron measured in terms of κ^{-1}, varies as β, in contrast to $\varepsilon \approx \beta^{3/2}$.

Therefore, it is not a simple matter to try to improve the Debye-Hückel distribution function. In this section, we shall limit our consideration to classical cases and shall examine more systematically the graphs we have adopted, in order to arrive at an improved result.

Generally, the contribution of a connected graph to $\rho_2(r)$ can be estimated by the following rules (*3*): (1) Each chain of graphs gives rise to a factor ε/κ^3.



(I)

(II)

Fig. 11.5. The classical diagrams of the order ε^2. Each solid line represents an effective interaction line.

(2) Each component of such a chain yields a factor n. (3) Each independent momentum variable representing an interaction contributes a factor κ^3.

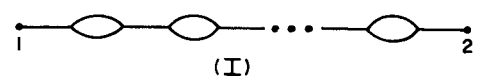
Since a series connection of chain diagrams yields only a chain diagram, corrections to the Debye–Hückel distribution will be obtained from diagrams which have at least one component graph with three interaction lines. Figure 11.5 illustrates the two simplest types of diagram which enter our consideration in this respect. Note that both diagrams have two independent momentum variables $\mathbf{q}$ and $\mathbf{q}'$. Applying the above rules, it is easy to find that both graphs contribute terms of order ε^2 to the correlation function. Explicitly, their contributions are given by (*3, 4*).

$$g^{(\mathrm{I})}(r) = -\frac{1}{n^2}\left(\frac{\kappa}{2\pi}\right)^6 \int [(1+s^2)(1+s'^2)(1+(\mathbf{s}-\mathbf{s}')^2]^{-1} \times \exp[i(\mathbf{s}-\mathbf{s}')\cdot\mathbf{x}]\, d\mathbf{s}\, d\mathbf{s}'$$
$$= -\frac{\varepsilon^2}{2\pi^2}\int \frac{1}{(1+q^2)}\frac{\tan^{-1}(q/2)}{q}\exp(i\mathbf{q}\cdot\mathbf{x})\, d\mathbf{q}, \tag{5.9}$$

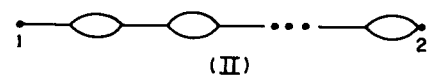
$$g^{(\mathrm{II})}(r) = \frac{1}{2n^2}\left(\frac{\kappa}{2\pi}\right)^6 \int \{(1+s^2)(1+s'^2)[1+(\mathbf{s}-\mathbf{s}')^2] \times [1+(\mathbf{s}'-\mathbf{s})^2]\}^{-1} \exp[i(\mathbf{s}-\mathbf{s}')\cdot\mathbf{x}]\, d\mathbf{s}\, d\mathbf{s}'$$
$$= \frac{\varepsilon^2}{4\pi^2}\int \frac{1}{(1+q^2)^2}\frac{\tan^{-1}(q/2)}{q}\exp(i\mathbf{q}\cdot\mathbf{x})\, d\mathbf{q}. \tag{5.10}$$

In going from the first line to the second line in these equations, we introduced a new set of variables by $\mathbf{Q} + (\mathbf{q}/2) = \mathbf{s}$ and $\mathbf{Q} - (\mathbf{q}/2) = \mathbf{s}'$, and we integrated over $\mathbf{Q}$.

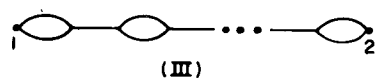
The diagrams in Fig. 11.5 can be used as units to construct complex chains as in Fig. 11.6. The reason for forming such chains is that the long-distance



(I)



(II)



(III)

Fig. 11.6. Composite chain diagrams.

behavior of the correlation function depends on the collective coupling of electrons. The type III chain can be shown to be of higher order. Corresponding to the first two types of complex chains, we find

$$g_c^{(I)}(x) = -\frac{\varepsilon}{ix\pi}\int_{-\infty}^{\infty} dq\, e^{iqx} q\left(q^2 + 1 + \frac{\varepsilon}{2q}\tan^{-1}\frac{q}{2}\right)^{-1}$$

$$= -\frac{\varepsilon}{x}e^{-\omega x}\left\{1 + \varepsilon\left[-\frac{1}{8}(\log 3) + \frac{1}{6}\right]\right\} + O(e^{-2x}), \tag{5.11}$$

and

$$g_c^{(II)}(x) = -\frac{\varepsilon^2}{2}\left[(\log 3)\frac{e^{-\omega x}}{x}\right] + O(e^{2x}). \tag{5.12}$$

Thus, adding these two contributions, we arrive at

$$g_c(x) = -\frac{\varepsilon e^{-\omega x}}{x}\left\{1 + \varepsilon\left[\frac{3}{8}(\log 3) + \frac{1}{6}\right] + O(\varepsilon^2)\right\}. \tag{5.13}$$

In these expressions,

$$\omega = 1 + \frac{\varepsilon}{8}(\log 3) + O(\varepsilon^2). \tag{5.14}$$

Equation (5.13) gives the first correction to the Debye distribution function. Finally, we can construct watermelon-type diagrams by using these complex chains and arrive at a form (Fig. 11.7):

$$g(x) = \exp\left[-A(\varepsilon)\frac{e^{-B(\varepsilon)}}{x}\right] - 1, \tag{5.15}$$

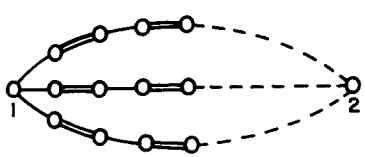


Fig. 11.7. A composite watermelon diagram.

which improves Eq. (5.8). The reader might want to try to find the two constants A and B.

11.6.* Charged Fermions

In this section, we consider a degenerate system of charged fermions. Our aim is to find their screening effects at low temperatures. The eigenvalues λ_j

are given by

$$\lambda_j(q) = \frac{1}{\hbar^3}\int d\mathbf{p}\, f^{(0)}(p)[1 - f^{(0)}(\mathbf{p}+\mathbf{q})]$$

$$\times \int_0^\beta \exp\frac{\alpha[p^2 - (\mathbf{p}+\mathbf{q})^2]}{2m}\exp\frac{2\pi ij\alpha}{\beta}\, d\alpha, \tag{6.1}$$

where $f^{(0)}(p)$ is the Fermi distribution. At low temperatures, $f^{(0)}(p)$ is

$$f^{(0)}(p) = \begin{cases} 0, & p > p_0, \\ 1, & p < p_0, \end{cases} \tag{6.2}$$

where p_0 is given by

$$p_0 = \hbar(3\pi^2 n)^{1/3}. \tag{6.3}$$

It is convenient to introduce dimensionless variables s, x, and y by

$$s = \frac{q}{p_0}, \qquad x = \frac{p_0 r}{\hbar}, \qquad y = \frac{4m\pi j}{\beta {p_0}^2}. \tag{6.4}$$

Recall that λ_j as a function of s and y is given by Eq. (8.9) of Chapter 10:

$$\lambda(s, y) = \left(\frac{2m}{\hbar^3}\right)\frac{p_0}{8\pi^2}F(s, y); \tag{6.5}$$

$$F(s, y) = 1 - \frac{1}{8}\left(s - \frac{4}{s} - \frac{y^2}{s^3}\right)\log\left|\frac{(s^2+2s)^2 + y^2}{(s^2-2s)^2 + y^2}\right|$$

$$- \frac{y}{2s}\left[\tan^{-1}\left(\frac{s^2+2s}{y}\right) - \tan^{-1}\left(\frac{s^2-2s}{y}\right)\right]. \tag{6.6}$$

The above eigenvalue expression must be multiplied by a factor 2 for electrons with spins. The summation over j of the eigenvalues may be replaced by integration over y at low temperatures. We arrive at the expression

$$\frac{\Delta\rho_2(r)}{n^2} - 1 = -\frac{3r_s}{2\pi^2}\left(\frac{3}{2\pi}\right)^{1/3}\frac{1}{x}$$

$$\times \int_0^\infty\int dy\, ds\, \frac{sF^2(s, y)\sin sx}{s^2 + {s_0}^2 F(s, y)}, \tag{6.7}$$

where

$$s_0 = \left(\frac{2\hbar}{a_0 p_0 \pi}\right)^{1/2} = \frac{\hbar\mu_0}{2^{1/2}p_0}, \qquad r_s = \left(\frac{4\pi}{3}{a_0}^3 n\right)^{-1/3}, \qquad {\mu_0}^2 = \frac{4p_0}{a_0\hbar}. \tag{6.8}$$

μ_0 is known as the *Thomas–Fermi screening constant*, and $\Delta\rho_2 = \rho_2(r) + I_2(r)$, $I_2(r)$ being given by Eq. (3.2). Here, a_0 is the Bohr radius.

The p.d.f. of fermions at 0°K with the exchange correction of Eq. (4.9) is illustrated schematically in Fig. 11.8 for several values of r_s. The dashed

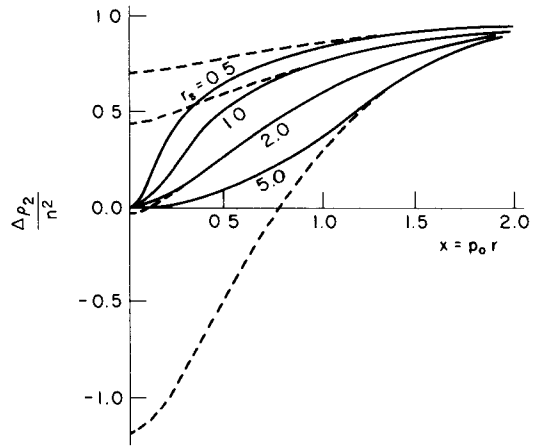


Fig. 11.8. Plot of $\Delta\rho_2(r)/n^2$ of a degenerate electron gas against a reduced distance $x = p_0 r/\hbar$. The dashed curves represent Eq. (6.7).

curves represent the chain diagram contribution given by Eq. (6.7) (*2*, *4*, *5*). When the density is larger, r_s^{-1} and also the kinetic energies of the particles are larger [see Eq. (6.3)]. Then, the effects of the Coulomb interactions become less important. This observation explains why the curves for smaller r_s are higher near the origin. As one sees, the results based on Eq. (6.7) are asymptotically correct at large distances.

In Fig. 11.8, the low-density curve, corresponding to $r_s = 5.0$, increases very rapidly. At distances around $x = 2.2$, it exceeds one and then decreases to one. The behavior of the p.d.f. of fermions near the characteristic distance $r \approx \hbar p_0^{-1}$ is expressed by

$$\frac{\Delta\rho_2(r)}{n^2} - 1 = -\frac{{p_0}^4}{4\pi^5 a_0 n^2}\left(1 + \frac{\hbar}{3\pi p_0 a_0}\right)\frac{e^{-\mu r}}{r}, \tag{6.9}$$

where μ is given by

$$\mu = \mu_0\left(1 + \frac{\hbar}{6p_0 a_0}\right). \tag{6.10}$$

Equation (6.9) gives a density correction to the Thomas–Fermi constant, just as Eq. (5.14) provides the Debye constant with a correction.

11.7.* Charged Bosons

As we have seen in the previous sections, the screening effects are due to the repulsive Coulomb forces. In fact, there is no screening if e, the charge, is zero. Due to the Coulomb repulsion, the correlation functions given by Eq. (5.6) and Eq. (6.9) are both negative. Comparing these expressions, we learn that Fermi statistics does not change the screening character of the correlation function. This result is understandable since Fermi statistics is repulsive in nature.

When Bose statistics is introduced, we expect some changes in the screening property of the correlation function because this amounts to having attractive pseudopotentials. Let us examine the case of charged bosons as a typical example (*4*, *5*).

The eigenvalues λ_j for degenerate bosons are given by Eq. (8.17) of Chapter 10:

$$\lambda_j(q) = 4mn\left(q^2 + \frac{4\pi jm}{\beta q^2}\right)^{-1}, \tag{7.1}$$

where n is the number density. Let us assume that the particles interact with each other with Coulomb forces. Using Eq. (7.1) in Eq. (4.8), we obtain

$$\rho_2(r) - n^2 = I_2(r) - \frac{4nA^4}{(2\pi)^3 r}\int_{-\infty}^{\infty} dx \times \int_0^{\infty} dy\, y^3[(y^4 + x^2)^2 + A^4(y^4 + x^2)]^{-1}\sin yr, \tag{7.2}$$

where $x = 4\pi jm/\beta$, $y = q/\hbar$, and

$$A^4 = \frac{2m}{\hbar^2}(8\pi ne^2), \tag{7.3}$$

e being the charge of the bosons.

The screening property of this system may be investigated by looking at the small-q or large-r behavior of $\rho_2(r)$. We first note that

$$\rho_2(r) - n^2 - I_2(r) = -\frac{2n}{(2\pi)^2 r}\int_0^{\infty} dy\, y^3[y^{-2} - (y^4 + A^4)^{-1/2}]\sin yr = \frac{n}{(2\pi)^2}\int_0^{\infty} dy\,[y^2 - (y^4 + A^4)^{1/2}]\cos yr, \tag{7.4}$$

and make use of the approximation

$$y^2 - (y^4 + A^4)^{1/2} \approx -\frac{2A^6}{y^4 + 2A^2y^2 + 2A^4}.$$

We obtain (4)

$$\frac{\Delta\rho_2(r)}{n^2} - 1 = -\frac{2^{1/2}r_s^{3/4}}{3^{1/4}}\, e^{-\nu x}(\cos bx + \nu \sin bx), \tag{7.5}$$

where

$$\Delta\rho_2(r) = \rho_2(r) - I_2(r),$$
$$2\nu^2 = (2^{1/2} + 1), \qquad 2b^2 = (2^{1/2} - 1), \qquad x = Ar.$$

The above result corresponds to the p.d.f. at 0°K. Therefore, the number density n may be replaced by that for the lowest energy state. On the other hand, if we do not let β be infinite, we arrive at an expression which includes the summation over the eigenvalues λ_j. The screening character for such a case differs from the above result. Indeed, we will have terms of the form $[e^{-\tau r}(\cos \tau r) - 1]/r$, where $\tau = 2^{-1/2}[(4\pi mj/\beta\hbar^2)^2 + A^4]^{1/4}$.

The asymptotic formula (7.5) shows oscillations. When it is plotted against x, the oscillations are independent of the density. Indeed, $\{[\Delta\rho_2(r)/n^2] - 1\}\, r_s^{-3/4}$ is a universal function of x for all densities.

11.8.* Phonon Spectrum and Spatial Correlations in a Hard-Sphere Bose Gas

According to Feynman, the excitation spectrum in liquid helium is given by

$$\varepsilon_k = \frac{\hbar^2k^2}{2mS(k)}, \tag{8.1}$$

where $S(k)$ is the structure factor, i.e., the Fourier transform of the spatial correlation function.

The phonon spectrum in liquid helium appears in the small-k region, which in the coordinate space corresponds to large distances. Phonons are due to collective modes of couplings of particles. At large distances, such modes are effectively represented by chain diagrams. Therefore, let us try to evaluate the p.d.f. of a hard-sphere Bose gas in the chain diagram approximation and see whether we can find the phonon spectrum.

A Bose gas condenses at low temperatures. The effect of particle correlations appears in the p.d.f. most strongly at short distances. As we approach the condensation temperature, the correlation effect extends to further and further distances.

To avoid complications due to condensation, let us consider a hard-sphere Bose gas at the lowest temperature. Since a hard-sphere potential does not have a Fourier transform, let us adopt an approximate pseudopotential (6):

$$\phi(r) = \frac{\hbar^2}{2m}[8\pi a\delta(\mathbf{r})], \tag{8.2a}$$

where a is the hard-sphere diameter and $\delta(\mathbf{r})$ is the Dirac delta function. The pseudopotential $\phi(r)$ has the Fourier transform

$$u(q) = \frac{\hbar^2(8\pi a)}{2m}. \tag{8.2b}$$

For liquid helium, the hard-sphere diameter a is approximately 2.56 Å. The volume per molecule, i.e., the inverse density, is $n^{-1} = 46.2$ Å^3. At the λ-point (2.18°K), the thermal de Broglie wavelength λ is 5.9 Å, and a/λ and $an^{1/3}$ are less than 1;

$$\frac{a}{\lambda} = 0.43, \qquad an^{1/3} = 0.71.$$

The pair distribution function in the chain diagram approximation is obtained by using Eq. (8.17) of Chapter 10 and Eq. (8.2b) in Eq. (4.8). The summation over j may be changed to integration as follows:

$$\sum_j \to \frac{\lambda}{2\pi^2}\int dx,$$

where $x = 2\pi^2 j/\lambda$, λ being the thermal de Broglie wavelength $(h^2/2\pi mkT)^{1/2}$.

Thus, the chain diagram approximation to the p.d.f. is expressed by (7–9)

$$\rho_2(r) = I_2(r) + n^2[1 - 4\pi^{-2}\gamma^{-1}aJ(r)], \tag{8.3}$$

where $I_2(r)$ is given by Eq. (3.2) of Chapter 10 and $\gamma = 4\pi an$, n being the number density. The integral $J(r)$ is defined by

$$J(r) = \mathrm{Re}\int [1 - y(y^2 + \gamma)^{-1/2}] \exp(2i\mathbf{y}\cdot\mathbf{r})\, d\mathbf{y} \tag{8.4}$$

where $\mathbf{y} = \mathbf{q}/\hbar$ and Re indicates real part. The integration may be performed in terms of the modified Bessel and the Struve functions.

At large distances, $I_2(r)$ is negligible. The asymptotic behavior of $J(r)$ may be obtained as follows. First, we note that

$$\begin{aligned} J(r) &= \mathrm{Re}\,\frac{2\pi}{ir}\int_0^\infty y[1 - y(y^2 + \gamma)^{-1/2}] \exp(2iyr)\, dy \\ &= \frac{2\pi}{r}\int_0^\gamma t^2(\gamma - t^2)^{-1/2} \exp(-2t\gamma)\, dt \\ &= \frac{2\pi\gamma}{r}\int_0^{\pi/2} (\sin^2\theta) \exp(-2\gamma^{1/2}r\sin\theta)\, d\theta. \end{aligned} \tag{8.5}$$

We then note that for large x the integral may be approximated by

$$J(r) = \frac{2\pi\gamma}{r}\int_0^\infty \theta^2 \exp(-2r^{1/2}r\theta)\, d\theta$$

$$\sim \frac{\pi}{2\gamma^{1/2}r^4}. \tag{8.6}$$

In other words, the p.d.f. decreases as r^{-4} (Fig. 11.9). This dependence corresponds to $S_k = (\hbar/2m)k/c$, which gives rise to a phonon spectrum

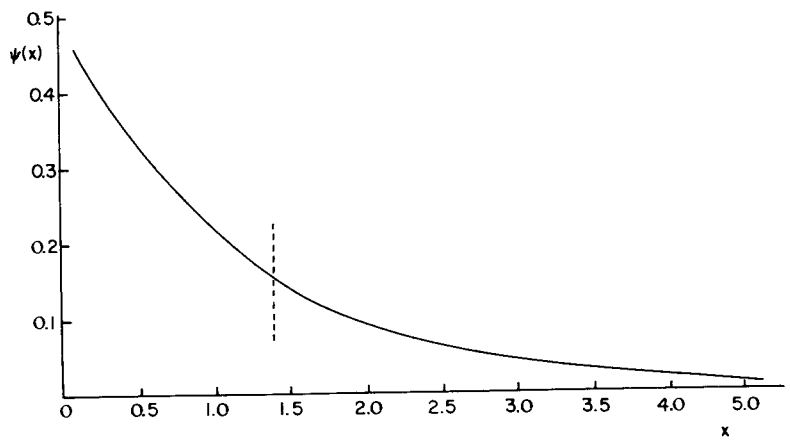


Fig. 11.9. The reduced correlation function $\psi(x) = [g(r) - 1](-4a/r)^{-1}$ of a hard-sphere Bose gas plotted against $x = 2\gamma^{1/2}r$. The curve becomes inaccurate for $x < 2^{1/2}$, as indicated by a dashed line, because exchange diagrams contribute at short distances.

$$\varepsilon = cp, \tag{8.7}$$

where c, the velocity of sound, is given by

$$c = \left(\frac{\hbar}{m}\right)\gamma^{1/2}. \tag{8.8}$$

Note that the phonon spectrum disappears in the absence of the hardcore because $\gamma = 4\pi an$.

The p.d.f. may be used directly for the evaluation of the ground-state energy. Indeed, we can make use of the formula

$$E_g = E_{\text{ideal}} + \tfrac{1}{2}V\int_0^1 d\xi \int dr\, \phi(r)\rho_2(r, \xi, a), \tag{8.9}$$

where ξ is a coupling constant which is artificially introduced to change the hard-sphere potential from zero to the actual value. Assuming that the ideal gas term is zero, we arrive at

$$E_g = \frac{\hbar^2}{2m}\left\{4\pi aN - \frac{1}{2}\sum_{\mathbf{k}}[k_a^2 + k^2 - k(k^2 + 2k_a^2)^{1/2}]\right\}, \tag{8.10}$$

where $k_a^2 = 8\pi an = 2\gamma$. The summation may be converted into an integration:

$$E_g = \frac{\hbar^2}{2m}\left\{4\pi anN + \frac{Vk_a^5}{4\pi^2}\int_0^\infty dy\, y^2\left[-1 - y^2 + y(y^2+2)^{1/2} + \frac{1}{2y^2}\right]\right\}.$$

Equation (8.10) involves a divergence due to the approximate pseudopotential. However, it can be removed correctly by adding a term $\frac{1}{2}(\hbar^2k_a^4/2k^2)$. The above integral includes this correction. We arrive at

$$E_g = \frac{\hbar^2}{2m}4\pi anN\left[1 + \frac{128}{15\pi^{1/2}}(a^3n)^{1/2}\right]. \tag{8.11}$$

The phonon excitation, if it is the only one in liquid He II at low temperatures, is important in explaining superfluidity. Suppose a macroscopic body M moves with velocity $\mathbf{v}$ in liquid helium. It will be slowed down if it generates excitations. From the conservation laws, we have

$$\tfrac{1}{2}Mv^2 = \tfrac{1}{2}Mv'^2 + \hbar\omega,$$
$$M\mathbf{v} = M\mathbf{v}' + \hbar\mathbf{k}.$$

These two equations yield

$$\mathbf{v}\cdot\mathbf{k} = \frac{(\hbar\mathbf{k})^2}{2M} + \omega,$$

which gives the lowest speed:

$$v_c = \frac{\hbar k}{2M} + \frac{\omega}{k}.$$

In the limit $M \to \infty$, this critical value is determined by ω/k. If ω is proportional to k, that is, if the phonon excitation is the only possible energy transfer, v_c is positive and finite. Thus, if the body moves with a velocity less than v_c, there will be no energy dissipation. Note that, if $\omega \sim k^2$, $v_c = 0$.

The above consideration is valid only for 0°K. At finite temperatures, inelastic scattering of phonons by the moving body should be taken into consideration. In actual He II, the critical velocity is much less than is predicted theoretically, indicating the possibility that other low-lying states are excited.

11.9.* Hard-Sphere Fermions

Contrary to some early predictions, the gas of ^{3}He was finally liquified in 1948 by Sydriak, Grilly, and Hammel at Los Alamos. Since then, considerable efforts have been made to find its thermodynamic properties, because then one would be able to establish (*i*) an independent proof of Sommerfeld's theory of the electronic specific heats of metals (see Fig. 4.12, p. 95), (*ii*) an independent confirmation of the essential roles played by statistics in the λ-transition in liquid ^{4}He, and (*iii*) the possible existence of "Cooper pairs," in analogy to superconductive electrons.

The experiments were very limited in the beginning because ^{3}He was difficult to obtain. From nuclear reaction of neutrons in lithium and the β-decay of the reaction product tritium, it has become possible to obtain sufficient quantities of ^{3}He.

The progress in the experiments on liquid ^{3}He has been remarkable. Around 1954, the lowest possible temperature was about 1°K. In five years, it went down to around 0.1°K. Now, it is somewhere around a few millidegrees, that is, 10^{-3} °K.

The ^{3}He particles obey Fermi statistics. Otherwise, the particles are similar to ^{4}He. Let us therefore investigate the properties of a hard-sphere Fermi gas by the methods analogous to those in the previous section.

The spatial correlation function of an ideal Fermi gas has been given in Eq. (1.14). If the particles have spins, it is expressed by (*10*)

$$\rho_2^{(0)}(r) - n^2 = -(2J+1)\lambda^{-6}G_{3/2}^2(z, r), \tag{9.1}$$

where J is the spin value and

$$G_{3/2}(z, r) = \frac{\lambda^3}{h^3}\int \frac{\exp(i\mathbf{p}\cdot\mathbf{r}/\hbar)}{z^{-1}\exp(\beta p^2/2m)+1}\,d\mathbf{p}$$

has been given in Eq. (2.9).

The evaluation of the pair distribution function may be made conveniently in momentum space. In analogy to Eq. (4.8), one can show that

$$\Delta\rho_2(r) - n^2 = (2\pi)^{-6}\int N(y)\exp(2i\mathbf{y}\cdot\mathbf{r})\,d\mathbf{y} \tag{9.2}$$

where $N(y)$ is given by

$$N(y) = \sum_{l=z}^{\infty}\frac{z^l u_l(y)}{(l-2)!}; \tag{9.3}$$

z is the absolute activity, $\mathbf{y} = \mathbf{q}/\hbar$, and $u_l(y)$, which represents the contribution from the connected graphs, is obtained from $U_l(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_l)$ introduced in Section 6.8 by integration over all the coordinates except the relative coordinate $\mathbf{r}_2 - \mathbf{r}_1$ which characterizes the correlation function.

To first order in the hard-sphere diameter a, two types of diagrams enter our consideration, one with and the other without an exchange between the two representative particles 1 and 2. We shall call the former type B and the latter type A.

For the A-type graphs, we find

$$N_A(y) = 8a(4\pi)^3 J(2J+1)G_{3/2}(z)\lambda^{-4} \times \sum_{n_i}\frac{(-z)^{n_1+n_2+n_3}}{(n_1+n_2+n_3)^{3/2}}\exp\left(-\frac{4n_2(n_1+n_3)}{n_1+n_2+n_3}\frac{\beta q^2}{2m}\right),$$

where λ is the thermal de Broglie wavelength and

$$G_{3/2}(z) = \frac{\lambda^3}{h^3}\int\frac{1}{z^{-1}\exp(\beta q^2/2m)+1}\,d\mathbf{q} = -\sum\frac{(-z)^l}{l^{3/2}}. \tag{9.4}$$

Also introducing

$$G_{1/2}(z) = \frac{\partial G_{3/2}(z)}{\partial(\log z)} = -\sum_l\frac{(-z)^l}{l^{1/2}}, \tag{9.5}$$

we arrive at the contribution to $\rho_2(r)$ from the A-type graphs:

$$\rho_2^{(A)}(r) = \frac{8aJ(2J+1)}{\lambda^7}G_{3/2}(z)G_{3/2}(z, r)G_{1/2}(z, r). \tag{9.6}$$

The function $G_\sigma(z)$ is related to $G_\sigma(z, r)$ of Eq. (2.9) as follows:

$$G_\sigma(z) = G_\sigma(z, 0). \tag{9.7}$$

At low temperatures, we can make use of the expansion

$$G_{3/2}(z) = \left(\frac{4}{3\pi^{1/2}}\right)(\log z)^{3/2} + \left(\frac{\pi^{3/2}}{6}\right)(\log z)^{-1/2} + O(\log z)^{-3/2}.$$

Thus, in the limit $\beta\to\infty$, we obtain

$$\rho_2^{(A)}(r) = \frac{2aJ(2J+1)}{3\pi^5}\frac{{k_0}^3}{r^4}(\sin k_0 r)(\sin k_0 r - k_0 r\cos k_0 r), \tag{9.8}$$

where $\hbar k_0$ is the Fermi momentum.

Similarly, the contribution from B-type diagrams is found to be given by

$$\rho_2^{(B)}(r) = -\frac{4aJ(2J+1)}{\lambda^7}\sum_{n_i}\frac{(-z)^l}{(n_1+n_2+1)^{3/2}(n_3+n_4+1)^{3/2}} \times\int_0^1\frac{\exp[-\pi r^2/4\beta F(x; n_i)]}{[F(z; n_i]^{3/2}}\,dx. \tag{9.9}$$

where

$$F(x; n_i) = (n_2 + x)\left(1 - \frac{n_2 + x}{n_1 + n_2 + 1}\right) + (n_4 + x)\left(1 - \frac{n_4 + x}{n_3 + n_4 + 1}\right). \quad (9.10)$$

Equation (9.9) yields the following expression in the limit $\beta \to \infty$:

$$\rho_2^{(B)}(r) = -\frac{4aJ(2J+1)}{\lambda^6 r} G_{3/2}^2(z, r)$$

$$= -\frac{4aJ(2J+1)}{4\pi^4}\frac{1}{r^7}(\sin k_0 r - k_0 r \cos k_0 r)^2. \quad (9.11)$$

At large distances, $\rho_2^{(B)}(r)$ is less important than $\rho_2^{(A)}(r)$. Thus, the particle correlation varies as

$$\rho_2^{(A)}(r) \propto a k_0^4 \frac{\sin 2k_0 r}{r^3}. \quad (9.12)$$

The oscillation is due to the sharpness of the Fermi distribution.

The equation of state can be obtained from $N(0)$ by simple z-integrations as follows:

$$\frac{pV}{kT} = \frac{V}{(4\pi)^3}\int^z dz \int^z dz\, z^{-2} N(0) \quad (9.13)$$

$$= -(2J+1)G_{5/2}(z) - 2J(2J+1)\frac{a}{\lambda}[G_{3/2}(z)]^2. \quad (9.14)$$

This yields the zero-point pressure

$$p_z = [(2J+1)(15\pi^2)^{-1} - 2J(2J+1)ak_0(9\pi^3)^{-1}]\frac{\hbar^2 k_0^5}{2m}. \quad (9.15)$$

On the other hand, the number density at 0°K is given by

$$n_0 = [(2J+1)(6\pi^2)^{-1} - 2J(2J+1)ak_0(3\pi^3)^{-1}]k_0^3. \quad (9.16)$$

Thus, we arrive at the ground-state energy per particle (*10, 11*)

$$\frac{E_g}{N} = \varepsilon_g = \frac{3}{5}\left(\frac{p_F^2}{2m}\right) + \frac{\hbar^2}{2m}(8\pi an)J(2J+1)^{-1}. \quad (9.17)$$

where $p_F = \hbar(6\pi^2 n/2J+1)^{1/3}$. The ground-state energy increases due to the hard-sphere repulsion. The contribution from the hard-sphere interaction to the ground-state energy is similar to that of bosons given by the first term of Eq. (8.11), but it is smaller. For instance, for $J = \frac{1}{2}$, the energy is one-half that for bosons.

From Eq. (9.17), one can easily evaluate the velocity of sound. As in the case of hard-sphere bosons, it is determined by the hard-sphere repulsion. However, when Eq. (9.17) is compared with Eq. (8.11), it is seen that the phonon spectrum is not the only low-lying excitation in the case of hard-sphere fermions. This leads to the absence of superfluidity in a hard-sphere fermion gas, in agreement with the experiments on liquid ^{3}He.

Our considerations of hard-sphere quantum gases in the last two sections were based on a perturbation method. It is a challenging problem to seek the exact solution of the Schrödinger equation for a delta-function-type interaction. For two and three dimensions, we expect a total collapse of a system with such interactions if the interactions are attractive. If they are repulsive, the problem is not very meaningful because the potentials are zero almost everywhere. However, for one dimension, the problem is not trivial. It has relations to some other problems in statistical mechanics. Since the treatment is rather specialized, we shall not go into details. However, we remark that the so-called Bethe Ansatz, concerning the symmetry of the wave function, has been used effectively in treating the case.

Problems

1. Show that the p.d.f for a pair of ideal fermions of parallel spins is given by

$$\rho_2^{\uparrow\uparrow}(r) = \frac{n^2}{2} - 2\left[\frac{1}{h^3}\int \frac{z\exp(-\beta p^2/2m)\exp(i\mathbf{p}\cdot\mathbf{r}/\hbar)}{z\exp(-\beta p^2/2m)+1}\,d\mathbf{p}\right]^2$$

and that for antiparallel spins is

$$\rho_2^{\uparrow\downarrow}(r) = \tfrac{1}{2}n^2,$$

where n, the number density, is given by

$$n = \frac{2}{h^3}\int \frac{z\exp(-\beta p^2/2m)}{z\exp(-\beta p^2/2m)+1}\,d\mathbf{p}.$$

Examine the behavior of $\rho_2^{\uparrow\uparrow}(r)$ at $r = 0$ and also

$$\rho_2(r) = \rho_2^{\uparrow\uparrow}(r) + \rho_2^{\uparrow\downarrow}(r).$$

2. At $T = 0$ ($\beta = \infty$), prove the following:

$$\rho_2^{\uparrow\uparrow}(r) = \frac{n^2}{2} - 2\left[\frac{1}{2\pi^2 r^3}(\sin k_0 r - k_0 r \cos k_0 r)\right]^2,$$

where $k_0 = (3\pi^2 n)^{1/3}$.

3. In the chain diagram approximation, find for fermions that

$$\rho_2^{\uparrow\uparrow}(r) = \frac{n^2}{2} - 2I_2(r) - \frac{1}{2}\sum_j \int \frac{u(q)[\lambda_j^2 - \lambda_j^2(q,r)]\exp(i\mathbf{q}\cdot\mathbf{r}/\hbar)}{1+\lambda_j u(q)}\, d\mathbf{q},$$

$$\rho_2^{\uparrow\downarrow}(r) = \frac{n^2}{2} - \frac{1}{2}\sum_j \int \frac{u(q)\lambda_j^2 \exp(i\mathbf{q}\cdot\mathbf{r}/\hbar)}{1+\lambda_j u(q)}\, d\mathbf{q},$$

where $\lambda_j(q)$ and $\lambda_j(q,r)$ are twice as large as those given by Eqs. (6.5) and (4.10).

4. Show that Eq. (7.20) of Chapter 10 may be expressed by

$$p\beta - p_0\beta = \frac{1}{h^3}\int_0^1 d\xi \int \sum_j \frac{[\lambda_j u(q)]^2\xi}{1+\lambda_j u(q)\xi}\, d\mathbf{q},$$

where ξ is a coupling constant.

5. The p.d.f. at short distances may be approximated for classical systems by

$$\rho_2(r) - n^2 = -\frac{1}{h^3\beta}\sum_j \int \lambda_j^2(q)u(q)\exp(i\hbar^{-1}\mathbf{q}\cdot\mathbf{r})\, d\mathbf{q}.$$

Evaluate the right-hand side for a classical electron gas, and discuss the behavior of the p.d.f. around $r = 0$.

6. The p.d.f. for a classical electron gas at large distances may be expressed by

$$\rho_2(r) = n^2 \exp[-\phi\ (r)] + nw_{31}(r) + n^2 w_{42}(r)$$

to order ε^2. Here, $\varepsilon = \beta\varepsilon^2\kappa$, $\phi_D = \varepsilon e^{-x}/x$, $x = \kappa r$, and $n^m w_s^m(r)$ represents the contribution of the order ε^{s-m} and n is the density. Find that w_{31} and w_{42} correspond to the graphs in Fig. 11.5.

7. In the above problem, show that [cf. Eq. (5.9)]

$$w_{31}(r) = \int [-\phi_D(r_{13})][-\phi_D(r_{32})]^2\, d\mathbf{r}_3$$

$$= \frac{\varepsilon^2}{2\pi nx}\,\mathrm{Re}\int_{-\infty}^{\infty} \frac{e^{ikx}}{k^2+1}\log\frac{2+ik}{2-ik}\, dk$$

$$= -\frac{\varepsilon^2}{2nx}[(\log 3)e^{-x} + e^{-x}Ei(-x) - e^{x}Ei(-3x)].$$

Here, Re indicates a real part, and $Ei(-x) = -\int_x^\infty (e^{-y}/y)\, dy$.

8. Similarly, show that [cf. Eq. (5.10)]

$$w_{42}(r) = \frac{\varepsilon^2}{2n^2x}\left[\frac{xe^{-x}}{4}(\log 3) - \frac{e^{-x}}{3} + \frac{e^{-x}}{4}(\log 3) + \frac{1}{3}e^{-2x}\right.$$

$$\left. + \frac{x}{4}e^{-x}Ei(-x) + \frac{x}{4}e^{x}Ei(-3x) + \frac{e^{-x}}{4}Ei(-x) - \frac{e^{x}}{4}Ei(-3x)\right].$$

9.* Show that, to the, order ε^2 the energy of a classical electron gas is given by

$$U = \tfrac{3}{2}NkT - \tfrac{1}{2}NkT[\varepsilon + \varepsilon^2(\gamma - \tfrac{3}{2} + \tfrac{1}{2}\log 3\varepsilon)],$$

where $\gamma = 0.57721$ is Euler's constant. [Use $\rho_2(r)$.]

GENERAL REFERENCE

A. Isihara, *Progr. Theoret. Phys. Suppl.* **44**, 1 (1969).

REFERENCES

1. S. Fujita, A. Isihara, and E. W. Montroll, *Bull. Cl. Sci. Acad. Roy. Belg.* **44**, 1018 (1958).
2. A. Isihara, *Phys. Rev.* **172**, 166 (1968).
3. A. Isihara, *Phys. Rev.* **178**, 412 (1969).
4. A. Isihara and M. Wadati, *Phys. Rev.* **183**, 312 (1969).
5. M. Wadati and A. Isihara, *Phys. Rev.* **B1**, 2156 (1970).
6. T. D. Lee, K. Huang, and C. N. Yang, *Phys. Rev.* **106**, 1135 (1957).
7. A. Isihara and D. D. H. Yee, *Phys. Rev.* **136**, A618 (1964).
8. A. Isihara and C. Nisteruk, *Phys. Rev.* **154**, 150 (1967).
9. A. Isihara and D. W. Jepsen, *Phys. Rev.* **158**, 112 (1967).
10. A. Isihara and A. K. Gupta, *Phys. Rev.* **160**, 225 (1967).
11. T. D. Lee and C. N. Yang, *Phys. Rev.* **113**, 1165 (1959).

12

Transport Phenomena in Degenerate Systems

In this chapter, we shall discuss the transport phenomena in quantum gases, using a quantum mechanical version of the classical Boltzmann equation. We shall assume the applicability of the Boltzmann equation and shall not try to verify the equation from first principles. At low temperatures, the applicability of the Boltzmann equation with the density-dependent terms has been questioned. Since the derivation of Boltzmann's equation for quantum systems has been based on various simplifying assumptions and yet requires involved mathematics, we shall avoid the discussion. More fundamental aspects of irreversible statistical mechanics will be considered in the next chapter.

12.1. Uehling–Uhlenbeck Equation

The Boltzmann equation for a classical gas may be modified to apply to a gas obeying quantum statistics in the following form (*1*, *2*):

$$\frac{\partial f}{\partial t} + \mathbf{v} \cdot \frac{\partial f}{\partial \mathbf{r}} + \mathbf{X} \cdot \frac{\partial f}{\partial \mathbf{v}} = -J(f),$$

$$J(f) = \iint [ff_1(1+\theta f')(1+\theta f_1') - f'f_1'(1+\theta f)(1+\theta f_1)] \times \mathbf{g} I(\vartheta, g)\, d\Omega\, d\phi. \tag{1.1}$$

This form of the collision term is due to Uehling and Uhlenbeck. Therefore, the quantum Boltzmann equation is sometimes called the *Uehling–Uhlenbeck equation.* The functions with primes correspond to "post collision," and

$$d\Omega = \sin\vartheta\, d\vartheta\, d\Phi,$$
$$d\phi = \left(\frac{m}{h}\right)^3 d\mathbf{v}. \tag{1.2}$$

In Eq, (1.1), $I(\vartheta, g)$ is the cross section for the relative velocity of magnitude g in the direction (ϑ, φ) for the collisions between particles of velocities $\mathbf{v}$ and $\mathbf{v}_1$. The term $J(f)$ reduces to the classical collision term if $\theta = 0$. We have $\theta = 1$ for Bose statistics and $\theta = -1$ for Fermi statistics. The appearance of the terms with θ is due to the quantum mechanical transition probability, which depends on the availability of the final state. At equilibrium, the integrand of Eq. (1.1) vanishes, and we obtain

$$f^{(0)} = \frac{1}{e^{\beta(\varepsilon-\mu)} - \theta}. \tag{1.3}$$

When the system deviates slightly from equilibrium, one can linearize Eq. (1.1) using the forms

$$f = f^{(0)}(1+\psi),$$
$$\psi = (1+\theta f^{(0)})\chi, \tag{1.4}$$

where ψ and χ are assumed to be small. The time derivative on the left-hand side may be converted to space derivatives by making use of the equations of continuity for the density, the momentum, and the energy:

$$\frac{d\rho}{dt} + \rho \operatorname{div} \mathbf{u} = 0,$$

$$\rho \frac{d\mathbf{u}}{dt} = \rho\mathbf{X} - \operatorname{grad} p, \tag{1.5}$$

$$\frac{d}{dt}\left(\frac{p}{\rho}\right) = -\frac{2}{3}\frac{p}{\rho} \operatorname{div} \mathbf{u},$$

where $\mathbf{u}$ is the average velocity, p is the pressure, and $\mathbf{X}$ is the acceleration due to an external force. Thus, we obtain

$$\frac{f^{(0)2}}{z\exp(-\xi^2)}\left[\frac{\boldsymbol{\xi}}{c^{1/2}T} \cdot (\operatorname{grad} T)\left(\xi^2 - \frac{5}{2}\frac{p}{nkT}\right) + \left(\boldsymbol{\xi}\boldsymbol{\xi} - \frac{\xi^2}{3}\mathbf{1}\right) : \operatorname{Sym} \nabla\mathbf{u}\right] = -J_0(\chi), \tag{1.6}$$

where

$$J_0(\chi) = \int f^{(0)} f_1^{(0)}(1 + \theta f^{(0)'})(1 + \theta f_1^{(0)'})\{\chi + \chi_1 - \chi' - \chi_1'\} g I(\vartheta, g)\, d\Omega\, d\phi \tag{1.7}$$

and

$$(\text{Sym } \nabla \mathbf{u})_{xy} = \frac{\partial u_x}{\partial y} + \frac{\partial u_y}{\partial x},$$
$$z = e^{\beta\mu}, \tag{1.8}$$
$$\xi = (\mathbf{v} - \mathbf{u})c^{1/2} = (\mathbf{v} - \mathbf{u})\left(\frac{m}{2kT}\right)^{1/2}$$

Equation (1.6) suggests the following form for the solution of the Boltzmann equation:

$$\chi = -\boldsymbol{\Lambda} \cdot \text{grad}(\log T) - \nabla \cdot \mathbf{u} \tag{1.9}$$

with

$$\xi f^{(0)2} \frac{\xi^2 - \frac{5}{2}(p/nkT)}{z[\exp(-\xi^2)]c^{1/2}} = J_0(\boldsymbol{\Lambda}),$$
$$\frac{2(\xi\xi - \frac{1}{3}\xi^2 \mathbf{1}) f^{(0)2}}{z \exp(-\xi^2)} = J_0(\boldsymbol{\Pi}). \tag{1.10}$$

Here, the unknown vector function $\boldsymbol{\Lambda}$ is related to the thermal conductivity because it appears as the coefficient to the temperature gradient. The unknown tensor function $\boldsymbol{\Pi}$ is related to the viscosity coefficient, as one can easily see. The essential point here is that the solution χ of the Uehling–Uhlenbeck equation is designed to be proportional to the temperature and velocity gradients. If χ has this proportionality, so has f.

12.1.1. Thermal Conductivity

We note that the thermal flux $\mathbf{q}$ is given by the average

$$\mathbf{q} = \int \tfrac{1}{2} f m(\mathbf{v} - \mathbf{u})^2 \mathbf{u}\, d\phi. \tag{1.11}$$

The flux $\mathbf{q}$ may be assumed to be proportional to the temperature gradient:

$$\mathbf{q} = -\kappa \frac{dT}{d\mathbf{r}}. \tag{1.12}$$

Using Eqs. (1.9), (1.10), and (1.4) in Eq. (1.11), we find

$$\frac{\mathbf{q}}{kT} - \frac{5}{2}\frac{p\langle(\mathbf{v} - \mathbf{u})\rangle}{kT} = \frac{1}{3}[\boldsymbol{\Lambda}, \boldsymbol{\Lambda}]\frac{\partial(\log T)}{\partial \mathbf{r}}, \tag{1.13}$$

where

$$[\boldsymbol{\Lambda}, \boldsymbol{\Lambda}] = \int \frac{f^{(0)2}\boldsymbol{\xi} \cdot \boldsymbol{\Lambda}}{zc^{1/2} \exp -\xi^2}\left(\xi^2 - \frac{5}{2}\frac{p}{nkT}\right) d\phi. \tag{1.14}$$

Thus, the conductivity is expressed by a vector function $\boldsymbol{\Lambda}$ in the form

$$\kappa = \tfrac{1}{3}[\boldsymbol{\Lambda}, \boldsymbol{\Lambda}], \tag{1.15}$$

which can be shown to be always positive.

12.1.2. Viscosity Coefficient

Similarly, the viscosity coefficient η is obtained from Eq. (3.9) of Chapter 6:

$$\sigma_{xx} = -p + \frac{2\eta}{3}\left(2\frac{\partial u_x}{\partial x} - \frac{\partial u_y}{\partial y} - \frac{\partial u_z}{z}\right) \tag{1.16}$$

in the form

$$\eta = \frac{kT}{10}[\boldsymbol{\Pi}, \boldsymbol{\Pi}], \tag{1.17}$$

where

$$[\boldsymbol{\Pi}, \boldsymbol{\Pi}] = \int \frac{f^{(0)2}\boldsymbol{\Pi}}{z \exp -\xi^2} : \left(\xi\xi - \frac{\xi^2}{3}\mathbf{1}\right) d\phi. \tag{1.18}$$

The bracketed quantity on the left side can be shown to be positive, just as the quantity in Eq. (1.14).

The results (1.15) and (1.17) are given in terms of unknown functions. We shall discuss the determination of these functions in the next section.

12.2. Transport Coefficients

The Uehling–Uhlenbeck equation is designed to yield for equilibrium the ideal quantum distribution functions instead of the Maxwell distribution function. As in the case of the classical Boltzmann equation, the derivation of the Uehling–Uhlenbeck equation has been tried with simplifying assumptions

such as the random-phase approximation and the Born approximation.

In this section, we shall determine the unknown vector and tensor functions $\mathbf{\Lambda}$ and $\mathbf{\Pi}$ which appeared in Eqs. (1.15) and (1.17). These functions may be expanded in complete sets of linearly independent functions $\mathbf{h}^{(r)}$ and $\mathbf{k}^{(r)}$ as follows:

$$\mathbf{\Lambda} = \sum_{r=1}^{\infty} c_r \mathbf{h}^{(r)}, \qquad \mathbf{\Pi} = \sum_{r=1}^{\infty} c_r \mathbf{k}^{(r)}, \tag{2.1}$$

where Eq. (1.6) of the previous section suggests the forms

$$\begin{aligned} \mathbf{h}^{(r)} &= \boldsymbol{\xi} H_r(\xi^2), \\ \mathbf{k}^{(r)} &= (\boldsymbol{\xi}\boldsymbol{\xi} - \tfrac{1}{3}\xi^2 \mathbf{1}) K_r(\xi^2). \end{aligned} \tag{2.2}$$

The unknown functions $H_r(\xi^2)$ are determined by the subsidiary conditions

$$\int \psi_k f^{(0)}(1 + \theta f^{(0)}) \mathbf{h}^{(r)} \, d\phi = 0, \qquad k = 1, \ldots, 5, \tag{2.3}$$

which state the conservation of ψ_k. Here, $\psi_1 = m$; $(\psi_2, \psi_3, \psi_4) = m\boldsymbol{\xi}$, $\psi_5 = m\xi^2$. These are the solutions of

$$J(\psi) = 0. \tag{2.4}$$

Thus, we find

$$H_r(\xi^2) = \xi^{2r} - \frac{\Gamma(r + \frac{5}{2})}{\Gamma(\frac{5}{2})} \frac{G_{r+3/2}}{G_{3/2}}. \tag{2.5}$$

Similarly,

$$K_r(\xi^2) = \tau^{2r-2},$$

where

$$G_s(\alpha) = \frac{1}{\Gamma(s)} \int_0^{\infty} \frac{x^{s-1}}{e^{x+\alpha} - \theta} \, dx. \tag{2.6}$$

This function has appeared in Eq. (4.13) of Chapter 4. We have $\theta = +1$ for bosons and $\theta = -1$ for fermions. To proceed further, we try to obtain the best results in replacing the infinite series in Eqs. (2.1) by finite series. For this purpose, we introduce the extremum conditions

$$[\mathbf{\Lambda} - \mathbf{\Lambda}^{(n)}, \mathbf{\Lambda} - \mathbf{\Lambda}^{(n)}] = 0, \qquad [\mathbf{\Pi} - \mathbf{\Pi}^{(n)}, \mathbf{\Pi} - \mathbf{\Pi}^{(n)}] = 0, \tag{2.7}$$

where $\mathbf{\Lambda}^{(n)}$ corresponds to truncating Eq. (2.1) after the nth term. Here,

$$\begin{aligned} [\mathbf{\Lambda}, \mathbf{A}] &= \int \frac{f^{(0)2} \boldsymbol{\xi} \cdot \mathbf{A}}{z c^{1/2} \exp -\xi^2} : \left(\xi^2 - \frac{5}{2} \frac{p}{nkT}\right) d\phi, \\ [\mathbf{\Pi}, \mathbf{B}] &= \int \frac{f^{(0)2} \mathbf{B}}{z \exp -\xi^2} : \left(\boldsymbol{\xi}\boldsymbol{\xi} - \frac{\xi^2}{3} \mathbf{1}\right) d\phi. \end{aligned} \tag{2.8}$$

We then obtain linear algebraic equations which can easily be solved. For instance, for $\mathbf{\Lambda}$, we can reduce one of the extremum conditions to

$$\left[\mathbf{h}^{(r)}, \mathbf{\Lambda} - \sum_{s=1}^{n} c_s \mathbf{h}^{(s)}\right] = 0, \tag{2.9}$$

which is of the form

$$\sum_{s=1}^{n} c_s a_{rs} - \alpha_r = 0. \tag{2.10}$$

Here,

$$a_{rs} = [\mathbf{h}^{(r)}, \mathbf{h}^{(s)}], \qquad \alpha_r = [\mathbf{\Lambda}, \mathbf{h}^{(r)}]. \tag{2.11}$$

For viscosity, we replace these coefficients by

$$b_{rs} = [\mathbf{k}^{(r)}, \mathbf{k}^{(s)}], \qquad \beta_r = [\mathbf{\Pi}, \mathbf{k}^{(r)}]. \tag{2.12}$$

In terms of these quantities, we can determine c_s, of Eqs. (2.1), and accordingly the conductivity and the viscosity coefficients. The results are

$$\begin{aligned} \kappa &= -k \frac{A_{\alpha\alpha}}{A}, \\ \eta &= -3kT \frac{B_{\beta\beta}}{B}. \end{aligned} \tag{2.13}$$

Here, $A = |a_{rs}|$; $B = |b_{rs}|$; and $A_{\alpha\alpha}$ and $B_{\beta\beta}$ are the determinants obtained from A and B by adding a final row and a column of elements α_r and β_r, $r = 1, 2, \ldots, n$, respectively, with $\alpha_{n+1,n+1} = 0$ and $\beta_{n+1,n+1} = 0$.

12.3. Transport Phenomena in Degenerate Systems

12.3.1. Fermi Gas

The application of the Uehling–Uhlenbeck equation to a degenerate Fermi gas was made by Tomonaga many years ago (*3*). He found that, at low temperatures, the viscosity and the thermal conductivity increase in proportion to T^{-2} and T^{-1}, respectively. These dependences, which are due to the sharpness of the Fermi distribution, are what we have obtained by an elementary method in Section 4.7. Later theories based on correlation functions (see Chapter 13) and a momentum-space consideration also confirmed the dependences (*4*, *5*). Although there are minor differences among the results, a typical result for the thermal conductivity is

$$\kappa \cdot \frac{7}{48(2^{1/2}\pi)} \frac{\varepsilon_0^{3/2}}{m^{1/2}} \frac{1}{T} \frac{1}{Q_1}, \tag{3.1}$$

and that for the viscosity is

$$\eta = \frac{4}{15(2^{1/2}\pi^3)} \frac{m^{1/2}\varepsilon_0^{5/2}}{(kT)^2} \frac{1}{Q_2}, \tag{3.2}$$

where ε_0 is the Fermi energy and Q_1 and Q_2 are effective cross sections.

For a hard-sphere model, the effective cross sections can be evaluated explicitly. If one wants to compare the theoretical results with the data on liquid ^{3}He, the hard-sphere diameter may be chosen as 1.16×10^{-8} cm. From the linear portion of the specific heat curve, the density and Fermi momentum are found to be approximately 0.082 gm cm^{-3} and 8.28×10^{-20} gm cm sec^{-1}, respectively. With this choice of parameters, one gets transport coefficients which are not very far from the experimental values for liquid ^{3}He.

12.3.2. Fermi Liquids

Applications of the above expressions (3.1) and (3.2) to liquid ^{3}He have to be questioned even if one can achieve some agreements, because the Uehling–Uhlenbeck equation is applicable only to dilute gases. Therefore, Abrikosov and Khalatnikov (*6*) developed a theory of transport phenomena in liquid ^{3}He based on the quasi-particle concept of Landau's. While their theory is interesting and gives the same temperature dependences as Eqs. (3.1) and (3.2), we shall discuss a different approach, that due to Rice (*7*), based on a consideration of scatterings of particles by spin fluctuations.[1]

Helium-3 atoms have spins. At low temperatures, liquid ^{3}He is nearly ferromagnetic but may have paramagnetic spin fluctuations. When quantized, these fluctuations become magnons. The scattering of particles from magnons in nearly ferromagnetic Fermi liquids can be used to explain equilibrium and nonequilibrium properties of liquid ^{3}He. In what follows, we consider collisions of bare fermions with magnons. The fermions flip their spins upon such collisions. Therefore, we introduce a singlet fermion distribution function $f_s(k)$, where s denotes two spin states, up and down. As we discussed in Section 8.11, the magnons obey Bose statistics. The distribution function of magnons of energy $\hbar\omega$ is

$$n(\omega) = [e^{\beta\hbar\omega} - 1]^{-1}. \tag{3.3}$$

The collision term for the time variation of $f(k)$ assumes the form

$$[1 - f_\uparrow(k)]f_\downarrow(\mathbf{k}+\mathbf{q})[n(\omega)+1] - f_\uparrow(k)[1 - f_\downarrow(\mathbf{k}+\mathbf{q})]n(\omega),$$

[1] A spin independent formalism is possible. See C. J. Pethick, *Phys. Rev.* **177**, 391 (1969).

if only the following processes are considered: (a) A fermion with momentum $\mathbf{k}$ and spin up absorbs a magnon with energy $\hbar\omega$ and momentum $\mathbf{q}$ and makes a transition to the state with momentum $\mathbf{k}+\mathbf{q}$ and spin down; (b) the inverse of (a).

The detailed balance condition for these processes is

$$[1 - f^{(0)}(\mathbf{k})]f^{(0)}(\mathbf{k}+\mathbf{q})[n(\omega)+1] = f^{(0)}(k)[1 - f^{(0)}(\mathbf{k}+\mathbf{q})]n(\omega),$$

where $\mathbf{k}$, $\mathbf{k}+\mathbf{q}$, and $\hbar\omega$ must satisfy the energy and momentum conservation laws, and where $f^{(0)}$ is the Fermi distribution function.

Because of the form of the collision term, it is convenient to express the deviation from equilibrium in the following form:

$$f_s(k) = f^{(0)}(k) - \frac{\partial f^{(0)}(k)}{\partial \varepsilon_k{}^0} \phi_s(k). \tag{3.4}$$

Here, $\phi_s(k)$ represents a function to be determined and $\varepsilon_k{}^0 = \hbar^2k^2/2m$ is the kinetic energy. On the right-hand side, $f^{(0)}$ is independent of the spin directions specified by s. Equation (3.4) assumes a slightly different form from the previous choice (1.4).

We arrive at

$$\frac{\partial f_\uparrow(k)}{\partial t} = -\int d\mathbf{k}'\, Q(\mathbf{k}\uparrow, \mathbf{k}'\downarrow)[\phi_\uparrow(k) - \phi_\downarrow(k')], \tag{3.5}$$

where the transition probability $Q(\mathbf{k}\uparrow, \mathbf{k}'\downarrow)$ is defined by

$$Q(\mathbf{k}\uparrow, \mathbf{k}'\downarrow) = \sum_{\mathbf{q}} \int d\omega\, A_q(\omega) n(\omega) f^{(0)}(k)[1 - f^{(0)}(k')] \times \delta(\varepsilon_{\mathbf{k}'}^0 - \varepsilon_{\mathbf{k}}{}^0 - \hbar\omega)\, \delta(\mathbf{k}' - \mathbf{k} - \mathbf{q}). \tag{3.6}$$

The $A_q(\omega)$ in the integrand is called the spectral density of the magnons.

Rice adopted a simple form for $A_q(\omega)$:

$$A_q(\omega) = \frac{\alpha\omega}{q}, \qquad 0 \leqslant |\omega| \leqslant \omega_q,$$
$$= 0, \qquad |\omega| > \omega_q, \tag{3.7}$$

where α is an adjustable constant and ω_q is a cutoff parameter. Using these two constants, he obtained the thermal conductivity in the following form:

$$\kappa^{-1} = aT - bT^2 + O(T^4), \tag{3.8}$$

where a and b are new constants.

In the first approximation, Eq. (3.8) gives the same temperature dependence of the thermal conductivity as Eq. (3.1). However, as the temperature increases, its second and higher terms give rise to deviations from the Fermi liquid behavior. These results are compared with experiment in Fig. 12.1, in

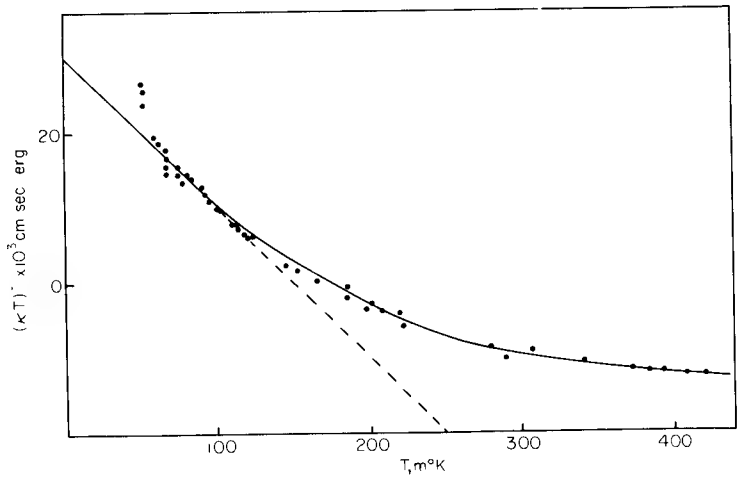


Fig. 12.1. The temperature dependence of the thermal conductivity of liquid ^{3}He at 7–10 cm Hg. The dots represent the data from Anderson *et al.* (*8*) and the solid curve represents Eq. (3.8).

which $(\kappa T)^{-1}$ is plotted against temperature (*7*). The solid curve is theoretical and the points represent the data of Anderson *et al.* (*8*). The broken line is the continuation of the straight line that was drawn through the points at ~ 100 m°K. The constant a is found to be 2.87×10^{-2} cgs, and the constant $b = 18 \times 10^{-2}$sec cm/erg °K.

Figure 12.2 illustrates the viscosity of liquid ^{3}He measured by Bettz and Wilks (*9*). It is seen that the T^{-2} dependence holds down to the lowest temperature, around 0.04°K. Below 0.07°K, the data are represented by an empirical formula

$$\eta = (3.8 \pm 0.6)T^{-2} \quad \mu\text{P}.$$

However, much lower values have been reported by other investigators. For instance, Abel *et al.* (*10*) reported $2.3T^{-2}$ μP. Thus, the experimental situation, though clear in outline, is still not quite definite in detail. The theory is also not quite definitely settled because several different types of theories lead to the same temperature dependences. Concerning deviations from Fermi liquid behavior, both Fermi gas and Fermi liquid theories would also yield correction terms, although detailed calculations have not been performed.

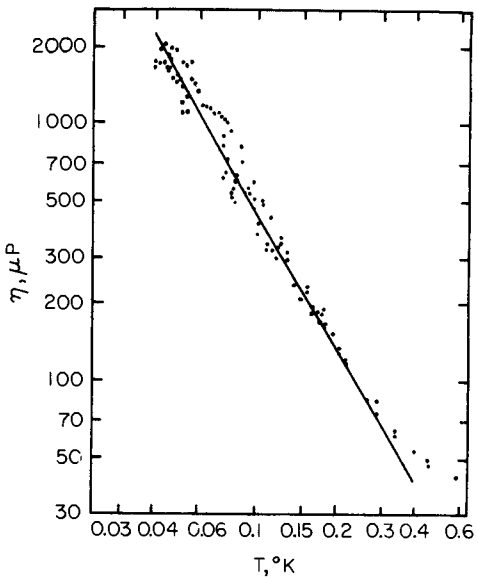


Fig. 12.2. The temperature dependence of the viscosity of liquid ^{3}He. [Data from Bettz and Wilks (*9*).]

12.3.3. Bose Gas

A degenerate Bose gas can be considered as a mixture of two kinds of particle, those in the lowest state and those in the excited states. Let the number density of the former be n_0 and that of the latter be n_1. We distinguish the collisions involving the latter kind of particle from those between the excited particles. Suppose two excited particles collide. After the collision, they may still be in excited states. The mean free path corresponding to this type of collision should be determined by the number density n_1. On the other hand, if one of the two particles drops into the zero-energy state, the mean free path should be dependent on n_0 and n_1. The same must be the case in the inverse collisions in which an excited particle collides with a zero-energy particle, and after the collision both are in excited states.

We assume that only excited particles can carry momenta and energies. Accordingly, only these particles contribute to the transport coefficients. If one adopts a hard-sphere model, the transport coefficients can be obtained explicitly. For instance, the viscosity coefficient is given by (*11*)

$$\eta = \frac{5G_{5/2}^2}{8a^2}\left(\frac{mkT}{\pi}\right)^{1/2} \frac{n_1}{16n_1 + G_{3/2}\, n_0}, \tag{3.9}$$

where a is the hard-sphere diameter, the functions $G_s = G_s(1)$ are defined by Eq. (2.6) and m is the mass of the particles.

Equation (3.9) does not involve any adjustable parameters. It shows that

the viscosity decreases with decreasing temperature because of the zero-energy particles. We have $G_{3/2} = 2.612$, and $n_0 + n_1 = n$ is the number density. Figure 12.3 illustrates Eq. (3.9). The hard-sphere diameter and mass are

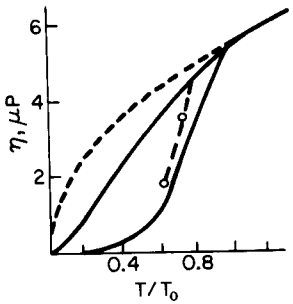


Fig. 12.3. The viscosity of a hard-sphere Bose gas. T_0 represents the temperature at which the Bose–Einstein condensation occurs. For comparison, the viscosity of He II is shown by a broken line.

2.10×10^{-8} cm and 3.2998×10^{-24} gm, respectively. The upper solid curve corresponds to $n_1/n = (T/T_0)^{3/2}$ and the lower solid curve corresponds to $n_1/n = (T/T_0)^{5.5}$, with T_0 the condensation temperature. The dashed curve represents the viscosity of the same gas without condensation. The circles represent the viscosity data for He II. These data are given only for comparing the order of magnitude rather than for a quantitative test. In fact, the condensation makes the applicability of the Uehling–Uhlenbeck equation difficult. The viscosity of liquid helium has also been analyzed in terms of elementary excitons.

12.4. Phonon–Phonon Scattering

The strange properties of liquid He II were attributed by Landau in 1941 to the energy spectrum, which was phononlike for small momenta and particlelike (rotons) for large momenta. Because the phonon spectrum is connected to the roton spectrum, a term which is nonlinear in momentum has to be added for higher momenta. Landau and Khalatnikov introduced

$$\varepsilon = cp(1 - \gamma p^2), \tag{4.1}$$

where γ is an adjustable parameter. Higher-order terms are neglected in Eq. (4.1). The cubic term in Eq. (4.1) causes scattering of phonons, by which transport phenomena in He II at low temperatures has been explained.

Corresponding to Eq. (4.1), let us split the hydrodynamic energy (Hamiltonian) of the liquid into a harmonic oscillator (phonon) part and a perturbation term (*12*). The total Hamiltonian of the system is given by a volume integration

$$H = \int [\tfrac{1}{2}\mathbf{v} \cdot \rho\mathbf{v} + \varepsilon(\rho)]\, d\mathbf{r}, \tag{4.2}$$

where ρ is the density, which determines the sound velocity. Since we are dealing with the quantum liquid, $\mathbf{v}$ should be considered to be an operator. The operator conjugate to $\mathbf{v}$ is the density, and we introduce the quantization condition

$$\rho(\mathbf{r}_1)\mathbf{v}(\mathbf{r}_2) - \mathbf{v}(\mathbf{r}_2)\rho(\mathbf{r}_1) = \frac{\hbar}{i} \nabla\delta(\mathbf{r}_1 - \mathbf{r}_2). \tag{4.3}$$

It is convenient to expand $\rho(\mathbf{r})$ and $\mathbf{v}(\mathbf{r})$ in Fourier series and transfer our treatment into momentum space:

$$\begin{aligned} \rho(\mathbf{r}) &= \rho_0 + \frac{1}{V^{1/2}} \sum_{\mathbf{k}} (\rho_{\mathbf{k}} \exp i\mathbf{k} \cdot \mathbf{r} + \rho_{\mathbf{k}}^* \exp -i\mathbf{k} \cdot \mathbf{r}), \\ \mathbf{v}(\mathbf{r}) &= \frac{1}{V^{1/2}} \sum_{\mathbf{k}} (\mathbf{v}_{\mathbf{k}} \exp i\mathbf{k} \cdot \mathbf{r} + \mathbf{v}_{\mathbf{k}}^* \exp -i\mathbf{k} \cdot \mathbf{r}). \end{aligned} \tag{4.4}$$

The quantization condition is then transformed into those for $\rho_{\mathbf{k}}$ ant $\mathbf{v}_{\mathbf{k}}$. The Hamiltonian can be decomposed into

$$\begin{aligned} H &= H_0 + H_1 + H_2, \\ H_0 &= \frac{c^2}{\rho_0} \sum (\rho_{\mathbf{k}} \rho_{\mathbf{k}}^* + \rho_{\mathbf{k}}^* \rho_{\mathbf{k}}), \\ H_1 &= \frac{\mathbf{v} \cdot \rho' \mathbf{v}}{2} + \frac{1}{3!} \frac{\partial}{\partial \rho} \left(\frac{c^2}{\rho}\right) \rho'^3, \\ H_2 &= \frac{1}{4!} \frac{\partial^2}{\partial \rho^2} \left(\frac{c^2}{\rho}\right) \rho'^4, \end{aligned} \tag{4.5}$$

where H_0 corresponds to the Hamiltonian of harmonic oscillators (phonons), and c is the sound velocity, ρ_0 is the equilibrium density and ρ' represents the deviation of the density from the equilibrium value.

The transition probability for phonon-phonon scattering is obtained in second-order perturbation theory from H_1 and in first-order from H_2 :

$$\begin{aligned} \langle \mathbf{p}, \mathbf{p}_1 | H | \mathbf{p}', \mathbf{p}_1' \rangle = \sum \frac{\langle \mathbf{p}, \mathbf{p}_1 | H_1 | \mathbf{q} \rangle \langle \mathbf{q} | H_1 | \mathbf{p}', \mathbf{p}_1' \rangle}{\varepsilon(\mathbf{p}) + \varepsilon(\mathbf{p}_1) - \varepsilon(\mathbf{q})} \\ + \langle \mathbf{p}, \mathbf{p}_1 | H_2 | \mathbf{p}', \mathbf{p}_1' \rangle. \end{aligned} \tag{4.6}$$

The calculation of the transition probability is not simple. Landau and Khalatnikov (*12*) found that the case $p \ll p_1$, i.e., the case when one of the

phonons has much lower momentum than the other, is important for collision processes at low temperatures. In such a case, one finds the cross section for phonon–phonon scattering averaged over all angles in the form

$$\sigma(p, p_1) = \frac{\pi(u+2)^4 {p_1}^4}{(96\pi\hbar^2\rho_0 c)^2\gamma}, \qquad p \ll p_1. \tag{4.7}$$

Note that γ, the dispersion in the phonon spectrum, appears in this expression. The quantity u is given by

$$u = \frac{\rho}{c^2}\frac{\partial c^2}{\partial \rho}. \tag{4.8}$$

The p^4 proportionality in Eq. (4.7) is analogous to the λ^{-4} proportionality in Rayleigh–Einstein scattering theory of light by density fluctuations.

For nonequilibrium, the distribution function of the elementary excitations may be determined from the equation

$$\frac{\partial f}{\partial t} + \frac{\partial f}{\partial \mathbf{r}}\cdot\frac{\partial H}{\partial \mathbf{p}} - \frac{\partial f}{\partial \mathbf{p}}\cdot\frac{\partial H}{\partial \mathbf{r}} = I(f), \tag{4.9}$$

$$I(f) = -\int c\sigma(\mathbf{p}, \mathbf{p}_1)[ff_1(1+f')(1+f_1') - f'f_1'(1+f)(1+f_1)]\frac{d\mathbf{p}}{h^3}, \tag{4.10}$$

where H is the Hamiltonian.

The collision term may be simplified by approximating f_1, f', and f_1' by the equilibrium functions. Expressing the deviation of f from equilibrium by δf, we obtain

$$I(f) = -\delta f \int c\sigma(\mathbf{p}, \mathbf{p}_1) f^{(0)-1} f^{(0)'} f_1^{(0)'}(1 + f_1^{(0)})\frac{d\mathbf{p}_1}{h^3}. \tag{4.11}$$

Therefore, a relaxation time τ may be defined by

$$\tau^{-1} = \int c\sigma(\mathbf{p}, \mathbf{p}_1) f^{(0)-1} f^{(0)'} f_1^{(0)'}(1 + f_1^{(0)})\frac{d\mathbf{p}_1}{h^3}. \tag{4.12}$$

The evaluation of τ based on Eq. (4.7) and the equilibrium distribution function

$$f^{(0)} = \frac{1}{e^{\beta\varepsilon} - 1} \tag{4.13}$$

is still not simple. However, for very low energy, the distribution function may be replaced by $(\beta\varepsilon)^{-1} = (\beta cp)^{-1}$. We then find

$$\tau^{-1} = \frac{5}{16}\frac{(u+2)^4}{(\hbar^2\rho_0)^2 c\gamma h^3}\left(\frac{kT}{c}\right)^6 p. \tag{4.14}$$

The T^6 dependence may be understood from Eq. (4.7) and Eq. (4.12) because $\sigma \propto {p_1}^4, p_1 \propto T$, and because $f^{(0)}$ and $1 + f^{(0)}$ are both proportional to T.

The differential cross section given by Eq. (4.7) is independent of the momentum p of phonons of small energy. Since p_1 can be considered to be proportional to T, the cross section varies as

$$\sigma(p, p_1) \propto T^4.$$

From a more detailed examination of the cross section based on Eq. (4.7), it is found that small-angle scatterings of phonons are most important. This means that a rapid exchange of energy between phonons takes place. In other words, phonons are nearly in equilibrium. Nevertheless, at extremely low temperatures, the time required for the establishment of energy equilibrium in a phonon gas is an important factor in determining transport properties.

At temperatures around 0.8°K, phonon–roton scattering must be taken into consideration. As in the case of phonon–phonon scattering, this is induced by perturbations of roton energy due to phonons.

At still higher temperatures, roton–roton scattering becomes important. These scatterings can be treated like those between hard-sphere particles. From phonon-roton scattering, Landau and Khalatnikov (*12*) found

$$\eta_p \begin{cases} = 8.7 \times 10^{-9} T^{-1/2} e^{\Delta/kT}, & T > 1°\text{K}, \\ = 7.8 \times 10^{-10}(T^{1/2} e^{-\Delta/kT} + 4.8 \times 10^{-6} T^5)^{-1}, & T < 0.8°\text{K}. \end{cases} \tag{4.15}$$

where the numerical values have been determined by taking experimental data into consideration.

Equation (4.15) shows that the phonon component of the viscosity coefficient increases as temperature decreases. Below 0.7°K, the dependence changes gradually to T^{-5}. On the other hand, above 1°K, the contribution from roton–roton scattering to the viscosity coefficient is estimated as

$$\eta_r = 1 \times 10^{-5} \text{ P}. \tag{4.16}$$

Finally, the viscosity of the normal component of liquid He II is

$$\eta_n = \eta_p + \eta_r. \tag{4.17}$$

Figure 12.4 illustrates the viscosity of the normal component of He II. The solid curve represents the theoretical result corresponding to Eq. (4.17). With parameters in the theory having been adjusted, the theoretical curve reproduces well the sharp increase of η_n at lower temperatures. However, the theory does not explain the viscosity increase with increasing temperature between 1.8°K and the λ-point.

So far, we have discussed only the viscosity coefficient. The thermal conductivity in He II is even more complicated. Experimentally, temperature

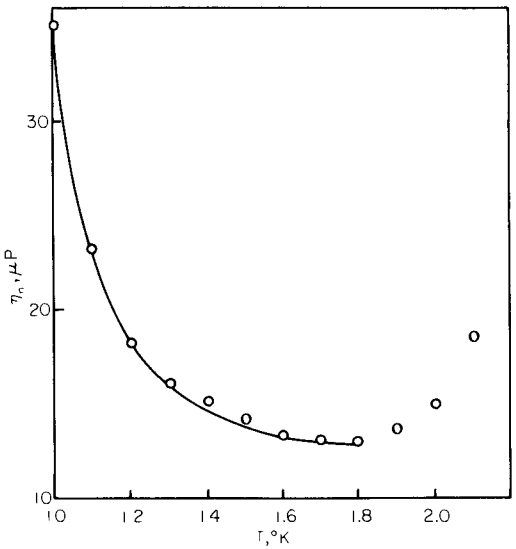


Fig. 12.4. The viscosity of the normal component of He II. The line represents the theoretical curve due to Landau and Khaltnikov (*12*). [Data from Bettz and Wilks (*9*).]

differences of only a few millidegrees produce large heat fluxes, and for high temperature gradients, the linear law of heat flow breaks down.

The magnitude of an effective thermal conductivity in a slit of width d may be expressed by

$$\kappa \propto 1.2 \times 10^5 T^{12.2} d^2.$$

For a slit of 10^{-3} cm width, the conductivity is approximately 10^{12} cal deg^{-1}/(cm sec) which is roughly 10^7 times as large as the conductivity of He I.

12.5. Electrical Conductivity of Metals

Let us denote by $f(k)$ the number density of the carriers of electricity in the state $\mathbf{k}$. The carriers obey Fermi–Dirac statistics and are scattered by scatterers. We shall let the scatterers play only a background role and neglect the electron–electron scattering. When the system is in a uniform electric field $\mathbf{E}$, the Boltzmann equation may be expressed as

$$-e\mathbf{E} \cdot \frac{\partial f}{\partial(\hbar\mathbf{k})} = \int [w(\mathbf{k}', \mathbf{k}) f'(1-f) - w(\mathbf{k}, \mathbf{k}') f(1-f')]\, d\mathbf{k}', \qquad (5.1)$$

where $w(\mathbf{k}, \mathbf{k}')$ represents the transition probability from the state $\mathbf{k}$ to $\mathbf{k}'$ and where $f' = f(k')$. The detailed balance condition for equilibrium requires

$$\begin{aligned} Q(\mathbf{k}', \mathbf{k}) &= w(\mathbf{k}, \mathbf{k}') f^{(0)}(1 - f^{(0)\prime}) \\ &= w(\mathbf{k}', \mathbf{k}) f^{(0)\prime}(1 - f^{(0)}) = Q(\mathbf{k}, \mathbf{k}'), \end{aligned} \qquad (5.2)$$

where $f^{(0)}$ is the Fermi distribution function, and where Q represents an effective transition probability. For small deviations from equilibrium, we use

$$f = f^{(0)}[1 + (1 - f^{(0)})\chi] \qquad (5.3)$$

to linearize Eq. (5.1). Using $\varepsilon_k{}^0 = \hbar^2 k^2/2m$, we obtain

$$-e\mathbf{E} \cdot \frac{\partial f^{(0)}}{\partial \varepsilon_k{}^0} \frac{\hbar\mathbf{k}}{m} = \int [\chi(k') - \chi(k)] Q(\mathbf{k}, \mathbf{k}')\, d\mathbf{k}'. \qquad (5.4)$$

The right-hand side represents a collisional change of $\chi(k)$ and may be effectively expressed as $f^{(0)}(1 - f^{(0)})\chi(k)/\tau$, where τ is a relaxation time. Thus,

$$f^{(0)}(1 - f^{(0)})\chi(k) = -\frac{\hbar}{m} \frac{\partial f^{(0)}}{\partial \varepsilon_k{}^0} \mathbf{k} \cdot e\mathbf{E}\tau. \qquad (5.5)$$

The electrical conductivity σ is obtained from

$$\frac{\mathbf{J}}{V} = \boldsymbol{\sigma} \cdot \mathbf{E},$$

where the current vector $\mathbf{J}$ can be evaluated from a statistical average of $e(\hbar\mathbf{k}/m)$. Using Eq. (5.5) and Eq. (5.3), we find [Note, $V^{-1} \sum_{\mathbf{k}} \to (2\pi)^{-3} \int d\mathbf{k}$.]

$$\begin{aligned} \mathbf{J} &= \frac{1}{(2\pi)^3} \int e\left(\frac{\hbar\mathbf{k}}{m}\right) f^{(0)}(1 - f^{(0)})\chi\, d\mathbf{k} \\ &= \frac{1}{(2\pi)^3} \left(\frac{e\hbar}{m}\right)^2 \int \tau\left(-\frac{\partial f^{(0)}}{\partial \varepsilon_k{}^0}\right) \mathbf{k}\mathbf{k} \cdot \mathbf{E}\, d\mathbf{k}. \end{aligned} \qquad (5.6)$$

Introducing a factor 2 due to spins, we arrive at

$$\boldsymbol{\sigma} = \frac{1}{4\pi^3} \left(\frac{e\hbar}{m}\right)^2 \int \tau\left(-\frac{\partial f^{(0)}}{\partial \varepsilon_k{}^0}\right) \mathbf{k}\mathbf{k}\, d\mathbf{k}. \qquad (5.7)$$

At low temperatures, the derivative $-\partial f^{(0)}/\partial \varepsilon_k{}^0$ behaves like a δ-function that has a peak at the Fermi energy. Accordingly, the conductivity is determined by the integral near the Fermi surface and the relaxation time τ is obtained by considering the changes in the direction of the $\mathbf{k}$ vectors. Let us assume elastic scattering and introduce

$$Q(\mathbf{k}, \mathbf{k}')\, d\mathbf{k}' = \delta(\varepsilon_k{}^0 - \varepsilon_k^{0\prime}) I(\mathbf{k}, \mathbf{k}')\, d\omega', \qquad (5.8)$$

where $I(\mathbf{k}, \mathbf{k}')$ is the differential cross section for $\mathbf{k}' \to \mathbf{k}$ and ω' is the solid angle representing the direction of $\mathbf{k}'$. Introducing Eqs. (5.4) and (5.8) into Eq. (4.3) and assuming a δ-function character for $\partial f^{(0)}/\partial \varepsilon_k{}^0$, we find

$$\frac{1}{\tau} = \int (1 - \cos\theta) I(k_0, \theta)\, d\omega', \qquad (5.9)$$

where $|\mathbf{k}| = |\mathbf{k}'| = k_0$ is the Fermi momentum divided by $\hbar$.

The electrical resistance of metals is due to the scattering of electrons by phonons. If an electron of momentum **k** collides with a phonon of momentum **q**, two kinds of processes take place. First, in what is called the *N* (*normal*) *processes*, the phonon is annihilated and the electron is scattered into a new state **k′**. The reverse processes can, of course, occur. These processes are expressed in terms of a vector equation by

$$\mathbf{k} + \mathbf{q} = \mathbf{k}'.$$

In the *N*-processes, the angle of deflection of the electron depends on the wave vector **q**. Since phonons carry only low momenta, particularly at low temperatures, this type of scattering can be considered on the Fermi surface: The vector **k** moves from one part of the surface to another.

The second class of scattering processes is known as the *Umklapp processes*. In these processes, **k′** does not lie on the same Fermi surface as **k**, but belongs to the Fermi sphere of the next Brillouin zone (Fig. 12.5).

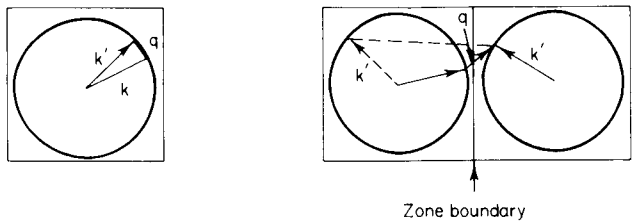


Fig. 12.5. Normal and Umklapp processes.

In other words, $\mathbf{k} + \mathbf{q}$ is beyond the Fermi surface, and only addition of a reciprocal lattice vector **g** would put **k′** on the original Fermi surface. Thus,

$$\mathbf{k} + \mathbf{q} + \mathbf{g} = \mathbf{k}'.$$

The electrical conductivity expression was obtained first by Grüneisen semiempirically, and then by Bloch theoretically based on the normal processes and on the Debye phonon spectrum. The Grüneisen–Bloch formula is

$$\sigma^{-1} = \frac{C}{M\vartheta}\left(\frac{T}{\vartheta}\right)^5 \int_0^{\vartheta/T} \frac{x^5\,dx}{(e^x - 1)(1 - e^{-x})}, \tag{5.10}$$

where M is the atomic weight, C is a constant, and ϑ is a characteristic temperature.

This formula gives the resistivity, which is proportional to T for high temperatures ($T \gg \vartheta$) and to T^5 for low temperatures. At low temperatures, phonons carry only very low momenta, and small-angle scattering becomes

important. According to Debye's theory, phonons exist at low temperatures in proportion to T^3. The scattering angle may be considered to be proportional to T/ϑ, since for small-angle scattering the effective cross section depends on the square of the scattering angle. In this way, we arrive at the T^5 law.

The Grüneisen–Bloch formula reproduces the general resistivity–temperature data of metals when C and ϑ are chosen suitably. However, it has been found that the Umklapp processes are much more important than the *N*-processes. According to Bailyn and Ziman, some 60–80% of the total resistivity of sodium is due to the Umklapp processes (*14*, *15*).

At low temperatures, the resistivity depends on impurity atoms. The typical examples are found in noble metals containing very small amounts of transition metals. Their resistivity shows either a minimum or both a minimum and a maximum. We shall discuss this phenomenon in the next section in more detail.

At very high temperatures, of the order of 1000°K, the resistivity of a large number of metals does not follow the linear proportionality law with respect to temperature. A possible mechanism for such a deviation has been given by Mott and others based on the two-band s-d scattering model.

Problem

Show that the thermal conductivity κ of metals is related to σ by

$$\kappa = \frac{\pi^2}{3}\frac{k^2}{e^2} T\sigma. \tag{5.11}$$

12.6.* Resistance Minima in Dilute Magnetic Alloys

Many years ago, Meissner and Voigt found a resistance minimum in supposedly pure conductors at low temperatures. Later, de Hass *et al.* (*16*), discovered the dependence of the minima on the purity of metals.

The phenomenon is most significant in dilute alloys of highly conductive metals, such as copper, silver, and gold including very small amounts (about 10^{-5} at. %) of magnetic impurity atoms such as chromium, manganese, or iron. Furthermore, Sarachik found, by measuring the susceptibility, that NbMo with a small amount of impurity iron atoms showed no resistance minimum when the sample did not have local magnetic moments.

Considering the magnetic moments of impurity atoms, it is clear that the interactions between the spins of conduction electrons and the moments play important roles in the phenomenon. Also, since the phenomenon takes place at low temperatures, Fermi statistics must be important. The resistance

is roughly proportional to the impurity concentration c, indicating that the phenomenon is due essentially to a single impurity atom interacting with the conduction electrons. The potential scattering of electrons by impurity atoms that does not involve spins may be considered unimportant for the phenomenon. Among those scatterings that flip the spins of impurity atoms, we consider simple processes such as:

(1) A conduction electron $\mathbf{k}^+$ (momentum $\mathbf{k}$ with a plus spin) is scattered to $\mathbf{q}'^-$ while changing the spin of the ith atom from M_i to $M_i + 1$. Then, it makes a transition to $\mathbf{k}'^+$, bringing $M_i + 1$ back to M_i, and it goes back to the original state $\mathbf{k}^+$. The matrix element for the transition probability for this type of scattering has the form

$$H_1(\mathbf{k}^+, M_i; \mathbf{q}'^-, M_i + 1)H_1(\mathbf{q}'^-, M_i + 1; \mathbf{k}'^+, M_i)H_1(\mathbf{k}'^+, \mathbf{k}^+) \frac{[1 - f^{(0)}(q')]}{(\varepsilon_k - \varepsilon_{q'})}.$$

(2) The inverse process may take place, in which the impurity atom loses one spin first and then recovers it later. The transition probability for this type of scattering has the form

$$-H_1(\mathbf{q}^-, M_i + 1; \mathbf{k}'^+, M_i)H_1(\mathbf{k}^+, M_i; \mathbf{q}^-, M_i + 1)H_1(\mathbf{k}'^+, \mathbf{k}^+) \frac{f^{(0)}(q)}{(\varepsilon_q - \varepsilon_{k'})}.$$

Here, H_1 represents the interaction Hamiltonian between the impurity and the conduction electron and $f^{(0)}$ is the Fermi distribution. The above probabilities must be summed over all the intermediate states and impurity atoms in the system. In addition, there are processes in which the electron $\mathbf{k}^+$ is scattered to the state $\mathbf{k}'^-$ instead of $\mathbf{k}'^+$ as in the above cases.

Considering all these possibilities, Kondo (*17*) obtained the relaxation time τ_k as follows:

$$\tau_k^{-1} = \frac{3zJ^2\pi S(S+1)c}{2\varepsilon_0 \hbar}[1 + 4Jg(\varepsilon_k)]. \tag{6.1}$$

Here, J is a coupling constant, z is the number of electrons per atom, ε_0 is the Fermi energy, S is the total spin, c is the concentration of impurity atoms, and

$$g(\varepsilon) = \frac{1}{N}\sum_{\mathbf{q}} \frac{f^{(0)}(q)}{\varepsilon_q - \varepsilon}, \tag{6.2}$$

where N is the total number of atoms.

The conductivity is given by

$$\sigma = -\frac{e^2\hbar^2}{12\pi^3 m^2}\int \tau_k k^2 \frac{df^{(0)}}{d\varepsilon_k^0} d\mathbf{k} \tag{6.3}$$

and the resistivity change $\Delta\rho$ due to the impurity atoms is

$$\Delta\rho = c\rho_M - c\rho_M \frac{\hbar^2 J}{\pi m k_0}\left[\int g(\varepsilon_k)\frac{df^{(0)}}{d\varepsilon_k} d\mathbf{k}\right], \tag{6.4}$$

where

$$\rho_M = 3\pi m J^2 S(S+1)\frac{V/N}{2e^2\hbar\varepsilon_0}, \tag{6.5}$$

V being the total volume.

The integral may be carried out to order $(kT/\varepsilon_0)^2$. We obtain

$$\Delta\rho = c\rho_M\left[1 + c\left(\frac{3zJ}{\varepsilon_0}\right)\log T\right]. \tag{6.6}$$

The total resistivity, which involves contributions from potential scattering due to lattice impurities, may be assumed to have the form

$$\rho = aT^5 + c\rho_0 - c\rho_1 \log T. \tag{6.7}$$

The constant a can be found from the pure crystal. Choosing the two arbitrary parameters ρ_1 and ρ_0 suitably and choosing $J = -0.022\varepsilon_0$, Kondo reproduced well the experimental data (Fig. 12.6).

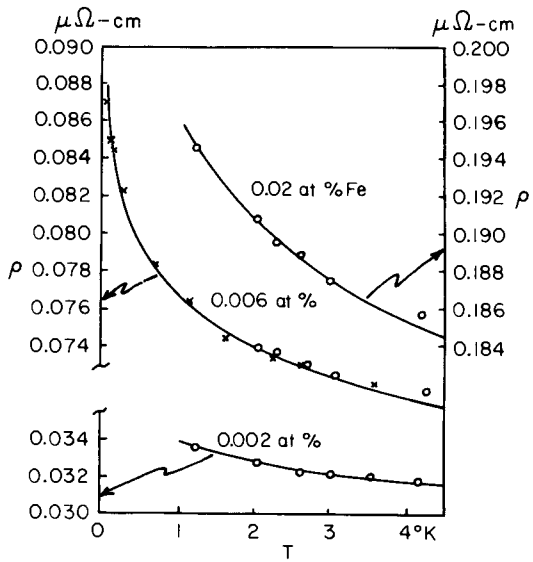


Fig. 12.6. The resistivity of gold with dilute iron impurity. The line represents the theoretical curve due to Kondo (*17*).

Concerning the sign of the coupling parameter J, it has been found that mixing of wave functions of conduction electrons and d-electrons leads to a negative sign for J. The agreement in Fig. 12.6 is due to the presence of the

logarithmic term in Eq. (5.7). Since this term causes divergence at 0°K, one must take higher-order perturbation terms into consideration as 0°K is approached. As in the case of an electron gas, the removal of the divergence can be made by summing an infinite series of the most divergent terms.

Problems

1. N particles are either in an excited state or in the ground state in accordance with

$$\frac{\partial n_1}{\partial t} = k_{01} n_0 - k_{10} n_1,$$

where n_1 and n_0 represent the fractions of particles in the two states. Assuming that the two states differ in the energy level by ε, discuss what will happen when the system deviates slightly from equilibrium due to oscillatory changes in temperature.

2. Discuss the basic differences between the elementary Maxwell theory of transport phenomena and the theories based on the Boltzmann equation.
3. Consider that the excitons in He II move with a small relative velocity $\mathbf{v}$ in a background of the superfluid which is at rest. The energy and momentum in the moving frame are related to the rest frame as follows:

$$\mathbf{p} = \mathbf{p}_0 + M\mathbf{v},$$

$$E = E_0 + \mathbf{v} \cdot \mathbf{p} + \tfrac{1}{2} M v^2,$$

where M is the total mass. Show that the average momentum current

$$\mathbf{j} = \int \mathbf{p} n(\varepsilon - \mathbf{v} \cdot \mathbf{p}) \frac{d\mathbf{p}}{h^3},$$

as seen in the moving system, is approximately equal to

$$\mathbf{j} = -\frac{1}{3} \mathbf{v} \int p^2 \frac{dn(\varepsilon)}{d\varepsilon} \frac{d\mathbf{p}}{h^3},$$

where $n(\varepsilon)$ is the number distribution.

4. In the above problem, if $\varepsilon = cp$ and if the distribution is given by the Bose distribution function, show that

$$\rho_n^{(p)} = \frac{2\pi^2 (kT)^4}{45 h^3 c^5},$$

where $\rho_n^{(p)}$ is the phonon contribution to the normal fluid density.

5. In Problem 3, if the distribution representing the roton part is

$$n = \exp \frac{-\Delta - (p - p_0)^2}{2\mu\, kT},$$

what result do you expect for $\rho_n^{(r)}$, which is the roton contribution to the normal fluid density?

6. Find, by a dimensional analysis, that the viscosity η of a plasma is proportional to $T^{5/2}$. Use the fact that η has the dimension of $ML^{-1}T^{-1}$ and that it is a function of the electronic charge e, the thermal energy kT, and the mass m. What about the case of the thermal conductivity?

REFERENCES

1. E. A. Uehling and G. E. Uhlenbeck, *Phys. Rev.* **43**, 552 (1933).
2. E. A. Uehling, *Phys. Rev.* **46**, 917 (1934).
3. S. Tomonaga, *Z. Physik.* **110**, 573 (1938).
4. M. Toda, *in* "Transport Processes in Statistical Mechanics" (I. Prigogine, ed.). Wiley (Interscience), New York, 1958.
5. H. Nishimura and H. Mori, *Progr. Theoret. Phys.* **26**, 967 (1961).
6. A. A. Abrikosov and I. M. Khalatnikov, *Rep. Progr Phys* **22**, 329 (1959).
7. M. J. Rice, *Phys. Rev.* **159**, 153 (1967).
8. A. C. Anderson, J. I. Connolly, O. E. Vilchs, and J. C. Wheatley, *Phys. Rev.* **147**, 86 (1966); W. R. Abel and J. C. Wheatley, *Phys. Rev. Lett.* **21**, 597 (1968).
9. D. S. Bettz and J. Wilks, *in* "Low Temperature Physics" (J. G. Daunt *et al.*, eds.). Plenum, New York, 1965.
10. W. R. Abel, A. C. Anderson, and J. C. Wheatley, *Phys. Rev. Lett.* **17**, 74 (1966).
11. A. Isihara, *J. Phys. Soc. Japan* **5**, 213, 216 (1950).
12. L. D. Landau and I. M. Khalatnikov, *Zh. Eksperim. i. Theo. Fiz.* **19**, 637 (1949).
13. P. P. Craig, W. E. Keller, and E. F. Hammel, *Ann. Physics* **21**, 72 (1963).
14. J. M. Ziman, *Proc. Roy. Soc. (London) Ser. A* **226**, 436 (1954).
15. M. Bailyn, *Phys. Rev.* **120**, 381 (1960).
16. W. J. de Haas, J. H. de Boer, and G. J. van den Berg, *Physica* **1**, 1115 (1933–1934).
17. J. Kondo, *Progr. Theoret. Phys.* **32**, 37 (1964); *Res. Electro. Lab. No.* 688 (1968).

13

Irreversibility and Transport Coefficients

In this chapter, we shall study the two major aspects of irreversible statistical mechanics: first, how transport coefficients can be obtained from a general consideration of the linear response of a system to external forces; and second, how the master equation can be derived from reversible dynamics.

For irreversible statistical mechanics, the year 1946 was made significant by the presentation of the BBGKY theories (1). In particular, Kirkwood (2) gave correlation function expressions for the transport coefficients of dense fluids and Bogoliubov (3) introduced two significant time scales, the time duration of an interaction and the relaxation time of the order of the time between collisions. These contributions greatly influenced subsequent developments (4).

In Chapter 7, we discussed the role of a correlation function for Brownian motion. The basic properties of response functions were studied by Callen and Welton (5, 6) around 1951. In 1955, Van Hove (7) derived the master equation quite generally by solving the Bloch equation for a weak coupling limit (8, 9). Further developments along this line have been achieved by Prigogine and his co-workers (10–13) and by many others.

For quantum systems, in 1956, Nakano (14) developed a linear response theory and derived a correlation function expression for the electrical conductivity. In the same year, Mori (15, 16) reported a similar correlation function formalism for viscosity. These and other important aspects of response functions were summarized by Kubo (17) in 1957. Also, since around 1957, the conductivity of metallic electrons has been studied from various viewpoints (18–21).

13.1. Response to External Forces

Let us consider a system described by a Hamiltonian H:

$$H = H_0 + H_1. \tag{1.1}$$

If the system is classical, the total distribution function is determined by the Liouville theorem:

$$\frac{\partial f}{\partial t} + [f, H] = 0, \tag{1.2}$$

where $[f, H]$ represents the Poisson bracket.

The system is assumed to be in equilibrium at $t = -\infty$. We consider that H_1 represents an external perturbation. To ensure the initial equilibrium situation, we switch on the perturbation at $t = -\infty$ adiabatically so that H_1 at $t = -\infty$ is zero. Furthermore, let us assume a form

$$H_1 = -A \cdot F(t), \tag{1.3}$$

with $F(-\infty) = 0$. Here, F represents a perturbing force and A is a corresponding "displacement." If A and F are vectors, the right-hand side is a scalar product.

The equilibrium distribution function f_0 satisfies

$$[f_0, H_0] = 0. \tag{1.4}$$

If f_1 represents that part of the distribution function that is due to the perturbation H_1,

$$f = f_0 + f_1, \tag{1.5}$$

we obtain

$$\frac{\partial f_1}{\partial t} + [f_1, H_0] = -[f, H_1]. \tag{1.6}$$

To solve Eq. (1.6), let us use the Green's function $g(t - t')$ satisfying

$$\frac{\partial g(t - t')}{\partial t} + [g(t - t'), H_0] = \delta(t - t'). \tag{1.7}$$

As we discussed in Section 7.5, the Green's function may be formally expressed as

$$g(t - t') = \exp iL(t - t'), \qquad t > t', \tag{1.8}$$

where the Liouville operator is defined such that

$$iLg + [g, H_0] = 0. \tag{1.9}$$

Then,

$$f_1(t) = -\int_{-\infty}^{t} [\exp iL(t-t')][f, H_1]_{t'}\, dt'. \tag{1.10}$$

Thus, adding f_0, we arrive at an integral equation

$$f(t) = f_0 - \int_{-\infty}^{t} [\exp iL(t-t')][f, H_1]_{t'}\, dt'. \tag{1.11}$$

The average value of a dynamical quantity B is obtained from

$$\langle B\rangle = \int Bf(t)\, d\Gamma, \tag{1.12}$$

where $d\Gamma$ represents the volume element of the phase space, and the distribution function f is to be normalized to 1. Thus, we obtain

$$\begin{aligned}\Delta\langle B\rangle &= -\int d\Gamma \int_{-\infty}^{t} \{\exp[i(t-t')L][A, f]_{t'}\} \cdot F(t')B\, dt' \\ &= -\int d\Gamma \int_{-\infty}^{t} [A, f]_{t'} B(t-t') \cdot F(t')\, dt',\end{aligned} \tag{1.13}$$

where $\Delta\langle B\rangle$ represents the change in $\langle B\rangle$ due to the perturbation, and where

$$B(t-t') = [\exp i(t-t')L]B. \tag{1.14}$$

A linear *response function* is defined by

$$\phi_{BA}(t) = \int d\Gamma\, [f_0, A]B(t) \tag{1.15a}$$

$$= \int d\Gamma\, f_0[A, B(t)], \tag{1.15b}$$

so that

$$\Delta\langle B(t)\rangle = \int_{-\infty}^{t} \phi_{BA}(t-t') \cdot F(t')\, dt'. \tag{1.16}$$

If the system obeys quantum statistics, the distribution function should be replaced by the density matrix that satisfies the Bloch equation

$$\frac{\partial\rho(t)}{\partial t} + \frac{1}{i\hbar}[\rho, H] = 0. \tag{1.17}$$

The solution to this differential equation with the H of Eq. (1.1) is obtained in the following form:

$$\begin{aligned}\rho(t) = \rho_0 - \frac{1}{i\hbar}\int_{-\infty}^{t} &\left[\exp\frac{-i(t-t')H_0}{\hbar}\right][A, \rho] \\ &\times \left[\exp\frac{i(t-t')H_0}{\hbar}\right] \cdot F(t')\, dt',\end{aligned} \tag{1.18}$$

where we have assumed $\rho(-\infty) = \rho_0$.

The linear response $\Delta\langle B(t)\rangle$ is given by

$$\begin{aligned}\Delta\langle B(t)\rangle &= \frac{1}{i\hbar}\,\mathrm{tr}\int_{-\infty}^{t}\left[\exp\frac{-i(t-t')H_0}{\hbar}\right][\rho_0, A] \\ &\quad\times\left[\exp\frac{i(t-t')H_0}{\hbar}\right]\cdot BF(t')\, dt' \\ &= \frac{1}{i\hbar}\,\mathrm{tr}\int_{-\infty}^{t}[\rho_0, A]B(t-t')\cdot F(t')\, dt' \\ &= \int_{-\infty}^{t}\phi_{BA}(t-t')\cdot F(t')\, dt',\end{aligned} \tag{1.19}$$

where

$$B(t) = e^{itH_0/\hbar}Be^{-itH_0/\hbar}, \tag{1.20}$$

and where the quantal response function $\phi_{BA}(t)$ is defined by

$$\phi_{BA}(t) = \frac{1}{i\hbar}\,\mathrm{tr}[\rho_0, A]B(t) \tag{1.21a}$$

$$= \frac{1}{i\hbar}\,\mathrm{tr}\,\rho_0[A, B(t)] \tag{1.21b}$$

$$= \frac{1}{i\hbar}\,\mathrm{tr}\,A[B(t), \rho_0]. \tag{1.21c}$$

The three expressions of Eq. (1.21) follow from the cyclic invariance of the trace.

Using the same cyclic invariant property of the trace and the commutability of ρ_0 with $e^{itH_0/\hbar}$, one can express $d\phi_{BA}/dt$ by the time derivative $\dot{A}$ given by

$$\dot{A} = \frac{i}{\hbar}(H_0A - AH_0). \tag{1.22}$$

We note that Eq. (1.21b) can be expressed by

$$\phi_{BA}(t) = \frac{1}{i\hbar}\,\mathrm{tr}\,\rho_0[A(-t), B]. \tag{1.23}$$

The time derivative $dA(-t)/dt$ is

$$\frac{dA(-t)}{dt} = -e^{itH_0/\hbar}\dot{A}e^{itH_0/\hbar}. \tag{1.24}$$

Transferring the time factor $e^{-itH_0/\hbar}$ in this expression to the right of B in Eq. (1.23), we obtain

$$\frac{d\phi_{BA}}{dt} = \frac{1}{i\hbar} \operatorname{tr} \rho_0[B(t), \dot{A}]. \tag{1.25}$$

13.2. Kramers–Kronig Relations for the Response Function

The expressions (1.21) for ϕ_{BA} are quite general. When A and B have components, the response function can be expressed similarly in terms of the components. For example, we have

$$\phi_{ij}(t) = \frac{1}{i\hbar} \operatorname{tr} \rho_0[A_j, B_i(t)], \tag{2.1}$$

with $B(t)$ given by Eq. (1.20) as the solution of

$$\frac{\partial B}{\partial t} - \frac{1}{i\hbar}[H_0, B] = 0. \tag{2.2}$$

The value $B_i(t)$ may be chosen to be $A_i(t)$, which represents a general extensive variable corresponding to an intensive generalized force $F(t)$.

Equation (2.1) gives immediately

$$\phi_{ij}(-t) = \frac{1}{i\hbar} \operatorname{tr} \rho_0[A_j(t), B_i]$$

if A_j and ρ_0 are invariant under time reversal.

We assume that $\phi_{ij}(t)$ has a Fourier transform

$$\chi_{ij}(\omega) = \int_0^\infty \phi_{ij}(t) e^{-i\omega t}\, dt. \tag{2.3}$$

In general, $\chi_{ij}(\omega)$ is complex. We shall introduce the real and the imaginary part response functions $\chi'(\omega)$ and $\chi''(\omega)$:

$$\chi_{ij}(\omega) = \chi'_{ij}(\omega) + i\chi''_{ij}(\omega), \tag{2.4}$$

where $\chi'(\omega)$ represents a reactive response and $\chi''(\omega)$ a dissipative response. We shall study the roles of these response functions in the next section.

In general, an analytic function $\chi(z)$ of a complex variable z may be expressed by Cauchy's integral as follows:

$$\chi(z) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi i} \frac{\chi(\omega)}{\omega - z} = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi i} \chi(\omega)\left(\frac{1}{\omega - z} \pm \frac{1}{\omega - z^*}\right), \tag{2.5}$$

if z is in the upper half-plane. Here, the last term does not contribute since $\chi(\omega)/[\omega - z^*]$ is analytic in the upper half ω-plane. Thus,

$$\chi(z) = \int_{-\infty}^{\infty} \frac{d\omega}{\pi i} \chi(\omega) \operatorname{Re} \frac{1}{\omega - z} \tag{2.6a}$$

$$= \int_{-\infty}^{\infty} \frac{d\omega}{\pi} \chi(\omega) \operatorname{Im} \frac{1}{\omega - z}. \tag{2.6b}$$

Taking the real and the imaginary parts of these two expressions, we obtain

$$\chi(z) = \int_{-\infty}^{\infty} \frac{d\omega}{\pi} [\operatorname{Im} \chi(\omega)]\left[\operatorname{Re} \frac{1}{\omega - z} + i \operatorname{Im} \frac{1}{\omega - z}\right]$$

$$= \int_{-\infty}^{\infty} \frac{d\omega}{i\pi} [\operatorname{Re} \chi(\omega)]\left[\operatorname{Re} \frac{1}{\omega - z} + i \operatorname{Im} \frac{1}{\omega - z}\right]. \tag{2.7}$$

Therefore, if z is in the upper half-plane,

$$\chi(z) = \int_{-\infty}^{\infty} \frac{d\omega}{\pi} \frac{\chi''(\omega)}{\omega - z}$$

$$= \int_{-\infty}^{\infty} \frac{d\omega}{i\pi} \frac{\chi'(\omega)}{\omega - z}. \tag{2.8}$$

We can use this formula for our χ_{ij}. If χ does not have singularities in the upper half-plane, we arrive at the Kramers–Kronig relations:

$$\chi'_{ij}(\omega) = P \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\chi''_{ij}(\omega')}{\omega' - \omega},$$

$$\chi''_{ij}(\omega) = -P \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\chi'_{ij}(\omega')}{\omega' - \omega}, \tag{2.9}$$

by making use of the identity

$$\frac{1}{x \mp i\varepsilon} = \frac{x \pm \varepsilon}{x^2 + \varepsilon^2}$$

$$= P\frac{1}{x} \pm \pi i \delta(x). \tag{2.10}$$

In these equations, P means principal value. For deriving Eqs. (2.9), a quantity ε may be introduced into the denominator of Eq. (2.8) and later equated to zero.

13.3. Symmetry Properties of Response Functions

In the previous section, the functions χ' and χ'' were introduced formally. We shall now show that χ'' plays an important role in our description of irreversibility.

If χ''_{ij} is defined by a Fourier transform of a commutator $[A_j, A_i(t)]$ for all t the antisymmetry relation

$$\chi''_{ij}(\omega) = -\chi''_{ji}(-\omega) \tag{3.1}$$

should hold. The function χ'' is defined in terms of A's which are to be Hermitian operators. Hence

$$[\chi''_{ij}(\omega)]^* = -\chi''_{ij}(-\omega). \tag{3.2}$$

Combining (3.1) and (3.2), we arrive at

$$[\chi''_{ij}(\omega)]^* = \chi''_{ji}(\omega). \tag{3.3}$$

In other words, χ'' is Hermitian.

If the Hamiltonian is invariant under time reversal, we find more generally

$$\chi''_{ij}(\omega) = -\varepsilon_j \varepsilon_i \chi''_{ji}(-\omega), \tag{3.4}$$

where ε is $+1$ for electric fields, while $\varepsilon = -1$ for magnetic fields. The relation (3.4) means that $\chi''(\omega)$ is odd in ω, and symmetric under the interchange of i and j if A_i and A_j have the same signature in ε.

For a system that is subject to a magnetic field H (or to some other perturbation that changes sign under time reversal), the magnetic susceptibility obeys the relation

$$\begin{aligned}\chi''_{ij}(\omega, H) &= -\chi''_{ij}(-\omega, H) \\ &= \chi''_{ji}(\omega, -H).\end{aligned} \tag{3.5}$$

These are examples of the *Onsager relations*.

The symmetry properties for χ' follows from Eq. (2.9):

$$\chi'_{ij}(\omega, H) = P \int_{-\infty}^{\infty} \frac{d\omega'}{\pi} \frac{\chi''_{ij}(\omega', H)}{\omega' - \omega}. \tag{3.6}$$

For instance, we find

$$\begin{aligned}\chi'_{ij}(\omega, H) &= \chi'_{ij}(-\omega, H) \\ &= \chi'_{ji}(\omega, -H).\end{aligned} \tag{3.7}$$

Essentially the same relations hold for the conductivity tensor.

We shall now prove that χ'' represents a dissipative response of a system to an external disturbance. For this purpose, we note that the average change

in energy is given by the rate of change of the expectation value of the Hamiltonian. Thus, if the perturbation Hamiltonian has the structure

$$H_1 = -\sum_i A_i F_i(t), \tag{3.8}$$

we find the rate of work

$$\begin{aligned}\frac{-d(\Delta W)}{dt} &= \sum_i \Delta \langle (A_i(t)) \rangle \dot{F}_i(t) \\ &= \sum_{ij} \int_{-\infty}^{t} \dot{F}_i(t) \phi_{ij}(t - t') F_j(t')\, dt'.\end{aligned} \tag{3.9}$$

Let us assume that the force F_i oscillates sinusoidally:

$$F_i(t) = \tfrac{1}{2}(F_i e^{-i\omega t} + F_i^* e^{i\omega t}). \tag{3.10}$$

By taking averages over a cycle, we find for the mean rate of change of energy

$$-\frac{d(\overline{\Delta W})}{dt} = \frac{i\omega}{4} \sum_{ij} [F_i^* \chi_{ij}(\omega) F_j - F_i \chi_{ij}(-\omega) F_j^*]. \tag{3.11}$$

Because of the antisymmetry and symmetry properties of χ''_{ij} and χ'_{ji} given by Eqs. (3.1) and (3.7), we obtain

$$\frac{d(\overline{\Delta W})}{dt} = \frac{\omega}{2} \sum_{ij} F_i^* \chi''_{ij} F_j, \tag{3.12}$$

which shows the role of χ'' in the energy dissipation.

13.4. Response in Canonical Ensembles

Let us consider the cases where equilibrium states are described by the canonical distribution

$$\rho_0 = \frac{e^{-\beta H_0}}{\operatorname{tr} e^{-\beta H_0}}. \tag{4.1}$$

The response function given by Eqs. (1.21) may be transformed into a convenient form. For this purpose, we make use of the identity (*17*)

$$\begin{aligned}[e^{-\beta H}, B] &= \int_0^{\beta} d\lambda\, e^{-\lambda H}[B, H] e^{\lambda H} e^{-\beta H} \\ &= \int_0^{\beta} d\lambda \frac{d}{d\lambda} (e^{\lambda H} B e^{-\lambda H}) e^{-\beta H} \\ &\quad i\hbar \int_0^{\beta} e^{-\beta H} \dot{B}(-i\hbar\lambda)\, d\lambda.\end{aligned} \tag{4.2}$$

The first identity can be proved by examining both sides for $\beta = 0$ and then comparing their derivatives with each other for $\beta \neq 0$. The purpose of the last expression is to have the derivative $\dot{B}$.

The identity (4.2) may be used for the operator $e^{-\beta H_0} = \rho_0$. Equations (1.21) become

$$\phi_{BA}(t) = \frac{1}{i\hbar} \operatorname{tr}[\rho_0, A]B(t)$$

$$= \int_0^\beta \operatorname{tr} \rho_0 \dot{A}(-i\hbar\lambda)B(t)\, d\lambda \tag{4.3a}$$

$$= -\frac{1}{i\hbar} \operatorname{tr}[\rho_0, B]A(-t)$$

$$= -\int_0^\beta \operatorname{tr} \rho_0 \dot{B}(-i\hbar\lambda)A(-t)\, d\lambda. \tag{4.3b}$$

Equations (4.3a) and (4.3b) represent the response function in the canonical ensemble. The two forms are equivalent, and the one that is more suitable to a given case can be chosen as required.

If $F(t)$ is constant from $t = -\infty$ to $t = 0$ and is zero for $t > 0$, the response $\Delta\langle B(t)\rangle$ of Eq. (1.19) is given by

$$\Delta\langle B(t)\rangle = F \int_t^\infty \phi_{BA}(t')\, dt'. \tag{4.4}$$

Therefore, it is convenient to introduce a function

$$\Phi_{BA}(t) = \int_t^\infty \phi_{BA}(t')\, dt', \tag{4.5}$$

which represents the relaxation of a physical quantity B. This function becomes, for the canonical distribution,

$$\Phi_{BA}(t) = \operatorname{tr} \int_0^\beta \rho_0 A(-i\hbar\lambda)B(t)\, d\lambda$$

$$- \lim_{t\to\infty} \operatorname{tr} \int_0^\beta \rho_0 A(-i\hbar\lambda)B(t)\, d\lambda$$

$$= \operatorname{tr} \int_0^\beta \rho_0 A(-i\hbar\lambda)B(t)\, d\lambda - \operatorname{tr} \int_0^\beta \rho_0 A_0 B_0\, d\lambda, \tag{4.6}$$

where A_0 and B_0 are the diagonal parts of A and B. The last term of Eq. (4.6) guarantees that $\Phi_{BA}(t)$ vanishes for $t \to \infty$. On the other hand, for $t = 0$,

$$\Phi_{BA}(0) = \chi_{BA}(0)$$

$$= \operatorname{tr} \int_0^\beta \rho_0 A(-i\hbar\lambda)B\, d\lambda - \beta \operatorname{tr} \rho_0 A_0 B_0$$

$$= \operatorname{tr} \int_0^\beta \rho_0[A(-i\hbar\lambda) - A_0](B - B_0)\, d\lambda. \tag{4.7}$$

In the classical limit, Eq. (4.6) reduces to

$$\Phi^0_{BA}(t) = \beta[\langle A(0)B(t)\rangle - \langle A_0 B_0\rangle]. \tag{4.8}$$

13.5. Transport Coefficients

13.5.1. Electrical Conductivity

Let us consider a system of charged particles in an electric field. The perturbation Hamiltonian may be chosen as

$$H_1(t) = -\sum_i e_i \mathbf{r}_i \cdot \mathbf{E}(t), \tag{5.1}$$

where the $\mathbf{r}_i$ are the coordinates of particles of charge e_i, and $\mathbf{E}(t)$ is given by

$$\mathbf{E}(t) = \mathbf{E}_0 e^{(i\omega + \varepsilon)t}, \tag{5.2}$$

where ε is a small parameter introduced in order that $\mathbf{E}(t) \to 0$ for $t \to -\infty$.

The conductivity σ can be determined from

$$\frac{\langle \mathbf{J}\rangle}{V} = \boldsymbol{\sigma} \cdot \mathbf{E}. \tag{5.3}$$

The current $\langle \mathbf{J}\rangle$ is given by taking the average of

$$\mathbf{J} = \sum_i e_i \dot{\mathbf{r}}_i = \sum_i \frac{e_i \mathbf{p}_i}{m_i}. \tag{5.4}$$

Using Eqs. (1.19) and (4.3a) for $A = \sum e_i \mathbf{r}_i$ and $B = \sum e_i \dot{\mathbf{r}}_i$, we arrive at a formula due to Nakano (*14*):

$$\boldsymbol{\sigma} = \operatorname{tr} \int_{-\infty}^{\blacksquare} \frac{1}{V} e^{i\omega(t'-t)} \int_0^\beta \rho_0 \mathbf{J}(-i\hbar\lambda)\, d\lambda\, \mathbf{J}(t - t')\, dt'. \tag{5.5}$$

The frequency dependent conductivity is obtained in the form

$$\boldsymbol{\sigma}(\omega) = \frac{1}{V} \operatorname{tr} \int_0^\infty e^{-(i\omega+\varepsilon)\tau}\, d\tau \int_0^\beta \rho_0 \mathbf{J}(-i\hbar\lambda)\mathbf{J}(\tau)\, d\lambda, \tag{5.6}$$

where ε is allowed to go to zero after the integral over τ is evaluated.

13.5.2. Diffusion Constant

Let us assume a gravitational or a centrifugal field in the z direction acting uniformly upon all the particles in the system (*22*). The Hamiltonian H_1 is

$$H_1 = -\sum_{i=1}^{N} z_i F. \tag{5.7}$$

The force F is related to the average velocity u by

$$F = u\zeta, \tag{5.8}$$

where ζ is the frictional constant. In turn, ζ is connected to the diffusion constant D through the Einstein relation

$$D = \frac{kT}{\zeta}. \tag{5.9}$$

Therefore, by evaluating the average of the velocities times kT per unit force, one obtains the diffusion constant:

$$D = \frac{kT}{F}\left\langle \frac{1}{m}\sum_{i=1}^{N} p_{z,i} \right\rangle = \frac{kT}{F}\langle J_z \rangle, \tag{5.10}$$

where p_{zi} is the z component of the momentum of the ith particle.

Thus, using Eq. (4.3a) for $A = \sum z_i$ and $B = J_z$, one finds

$$D = kT \operatorname{tr} \int_0^\infty d\tau \int_0^\infty d\lambda \rho_0 J_z(-i\hbar\lambda) J_z(\tau). \tag{5.11}$$

13.5.3. Viscosity Coefficient

For a homogeneous, Newtonian, incompressible fluid, the shear stress $\boldsymbol{\sigma}$ is proportional to the velocity gradient tensor $\mathbf{q}$ with the shear viscosity η as the proportionality constant [see Eq. (3.9) of Chapter 6]

$$\boldsymbol{\sigma} = 2\eta\mathbf{q}. \tag{5.12}$$

For a shear flow in the xy plane, the symmetrized velocity gradient tensor $\mathbf{q}$ can be given the components:

$$q_{xy} = q_{yx} = \frac{1}{2}\left(\frac{\partial u_y}{\partial x} + \frac{\partial u_x}{\partial y}\right), \qquad \text{all other } q_{ij} = 0, \tag{5.13}$$

where u_x and u_y are the x and y components of the average velocity of the system. The stress tensor is obtained by consideration of a momentum transfer and a drag effect due to potentials. We can write

$$\boldsymbol{\sigma} = \boldsymbol{\sigma}_K + \boldsymbol{\sigma}_\phi = \frac{\langle \mathbf{J}_K \rangle + \langle \mathbf{J}_\Phi \rangle}{V};$$

$$\mathbf{J}_K = -\sum_i m\left(\frac{\mathbf{p}_i}{m} - \mathbf{u}\right)\left(\frac{\mathbf{p}_i}{m} - \mathbf{u}\right); \tag{5.14}$$

$$\mathbf{J}_\Phi = \frac{1}{2}\sum_{j\neq i}\sum \left[\frac{(\mathbf{r}_j - \mathbf{r}_i)(\mathbf{r}_j - \mathbf{r}_i)}{r_{ji}}\frac{\partial \phi_{ji}}{\partial r_{ji}}\right].$$

Let us assume that the velocity gradient $\mathbf{q}$ oscillates sinusoidally. If $\mathbf{s}$ is the strain tensor, $\mathbf{q}$ can be given in the form

$$\mathbf{q} = \dot{\mathbf{s}}, \qquad \mathbf{s} = \mathbf{s}_0 e^{(i\omega+\varepsilon)t}, \tag{5.15}$$

where ε is a small, positive parameter to let the perturbation vanish at $t = -\infty$. We note that

$$\int_{-\infty}^{t} \boldsymbol{\sigma}\, dt = \int_{-\infty}^{t} 2\eta\dot{\mathbf{s}}\, dt = 2\eta\mathbf{s}$$

$$= \frac{1}{V}\left\langle \int_{-\infty}^{t} \mathbf{J}\, dt \right\rangle. \tag{5.16}$$

The last equality suggests the method of evaluating the viscosity coefficient η: We evaluate the average of a physical quantity $B = \int_{-\infty}^{t} J_{xy}\, dt$. The perturbation Hamiltonian for the system of volume V may be chosen to be

$$H_1 = -(-J_{xy})(2s_{xy}) = -AF, \tag{5.17}$$

with $A = -J_{xy}$ and $F = 2s_{xy}$. According to Eq. (4.3b), the response to this perturbation is given by

$$\phi_{BA}(t) = \int_0^\beta \operatorname{tr}[\rho_0 \mathbf{J}(-i\hbar\lambda)\mathbf{J}(-t)]\, d\lambda. \tag{5.18}$$

Thus, Eq. (1.19) becomes

$$\eta s_{xy} = \frac{1}{V}\int_{-\infty}^{t} \phi_{BA}(t-t')s_{xy}(t')\, dt'$$

$$= \frac{1}{V}\int_0^\infty \phi_{BA}(\tau)s_{xy}(t-\tau)\, d\tau$$

$$= \frac{1}{V}\left[\int_0^\infty \phi_{BA}(\tau)e^{-(i\omega+\varepsilon)\tau}\, d\tau\right]s_{xy}. \tag{5.19}$$

We arrive at

$$\eta(\omega) = \frac{1}{V}\lim_{\varepsilon\to 0}\int_0^\infty d\tau\, e^{-(i\omega+\varepsilon)\tau} \operatorname{tr}\int_0^\beta \rho_0 J_{xy}(-i\hbar\lambda)J_{xy}(-\tau)\, d\lambda. \tag{5.20}$$

The correlation function $J_{xy}(-i\hbar\lambda)J_{xy}(-\tau)$ can be written as $J_{xy}J_{xy}(-\tau+i\hbar\lambda)$. In component form, the static viscosity η is given by

$$\eta=\frac{1}{V}\int_0^\infty d\tau \operatorname{tr}\int_0^\beta \rho_0 J_{xy}(-i\hbar\lambda)J_{xy}(-\tau)\,d\lambda. \tag{5.21}$$

13.6. Reduction of Transport Coefficients

The transport coefficients such as given by Eqs. (5.6) and (5.11) can be expressed by time correlation functions of real variables. For instance, the conductivity is given for $V=1\ \text{cm}^3$ by

$$\sigma(\omega)=\lim_{\varepsilon\to 0}\int_0^\infty e^{-(i\omega+\varepsilon)\tau}\,d\tau\int_{-\infty}^\infty K(t-\tau)\langle \mathbf{J}(0)\mathbf{J}(t)\rangle\,dt, \tag{6.1}$$

where $\langle\cdots\rangle$ represents the average over a canonical ensemble.

In general, for

$$\Phi_{BA}(\tau)=\operatorname{tr}\int_0^\beta d\lambda\,\rho_0 A(-i\hbar\lambda)B(\tau) \tag{6.2}$$

we want to show that

$$\begin{aligned}\Phi_{BA}(t)&=\int_{-\infty}^\infty K(t-t')\Psi_{BA}(t')\,dt',\\ \Psi_{BA}(\tau)&=\operatorname{tr}\,\rho_0 A(0)B(\tau).\end{aligned} \tag{6.3}$$

For this purpose, let us choose $H_0=\hbar\omega$ and $\rho_0\propto e^{-\beta H_0}$, and define a Fourier transform of $\Phi_{BA}(\tau)$:

$$\tilde{\Phi}_{BA}(\omega)=\frac{1}{2\pi}\int_{-\infty}^\infty e^{-i\omega\tau}\Phi_{BA}(\tau)\,d\tau. \tag{6.4}$$

Introducing Eq. (6.2), using $\tau+i\hbar\lambda=t$ as a new variable and performing the λ integration first, we find the result

$$\tilde{\Phi}_{BA}(\omega)=\frac{1-e^{-\beta\hbar\omega}}{\hbar\omega}\frac{1}{2\pi}\int_{-\infty}^\infty dt\,e^{-i\omega t}\langle AB(t)\rangle. \tag{6.5}$$

Equation (6.5) leads to $\tilde{K}(\omega)$, the Fourier transform of the kernel $K(t-t')$:

$$\tilde{K}(\omega)=\frac{1-e^{-\beta\hbar\omega}}{2\pi\hbar\omega}. \tag{6.6}$$

In Eq. (6.5), we have modified the path of the t-integration from $(-\infty+i\hbar\lambda,\ \infty+i\hbar\lambda)$ to $(-\infty,\infty)$ under the assumption that $\operatorname{tr}\langle AB(t)\rangle$ is analytic in t in the domain $0\leqslant \operatorname{Im} t\leqslant\lambda\hbar$.

The integrand of Eq. (6.5), multiplied by $e^{-\beta\hbar\omega}$, can be written as

$$\begin{aligned}\operatorname{tr}[e^{-\beta\hbar\omega}A(0)e^{\beta\hbar\omega}e^{-\beta\hbar\omega}B(t)]&=\operatorname{tr}[e^{-\beta\hbar\omega}B(t)A(i\beta\hbar)]\\&=\operatorname{tr}[e^{-\beta\hbar\omega}B(t-i\beta\hbar)A].\end{aligned}$$

By using $(t-i\beta\hbar)$ as a new t variable, we find

$$\tilde{\Phi}_{BA}(\omega)=\frac{e^{\beta\hbar\omega}-1}{2\pi\omega\hbar}\int_{-\infty}^\infty e^{-i\omega t}\langle B(t)A\rangle\,dt. \tag{6.7}$$

Thus, combining Eqs. (6.5) and (6.7), we arrive at

$$\tilde{\Psi}_{BA}(\omega)+\tilde{\Psi}_{AB}(\omega)=\hbar\omega\coth(\tfrac{1}{2}\beta\omega\hbar)\tilde{\Phi}_{BA}(\omega), \tag{6.8}$$

where $\tilde{\Psi}_{BA}(\omega)$ is the Fourier transform

$$\tilde{\Psi}_{BA}(\omega)=\frac{1}{2\pi}\int e^{-i\omega t}\langle AB(t)\rangle\,dt. \tag{6.9}$$

At high temperatures, we can use

$$\tilde{K}(\omega)\to\tilde{K}_0(\omega)=\frac{\beta}{2\pi}. \tag{6.10}$$

This gives

$$K_0(t)=\beta\delta(t),\qquad \Phi^0_{BA}(t)=\Psi\beta_{BA}(t) \tag{6.11}$$

in agreement with Eq. (4.8).

We have demonstrated important roles played by time correlation functions for transport coefficients. However, these expressions are formal and need reductions to explicit forms.

Let us consider reducing the electrical conductivity expression (6.1) to a more explicit form for a simple case (*18–21*). First, we note that Eq. (6.11) yields

$$\sigma=\beta\frac{1}{V}\int_0^\infty\langle \mathbf{J}\mathbf{J}(t)\rangle\,dt \tag{6.12}$$

for high temperatures and the static case $\omega\to 0$. Of course, $\mathbf{J}(t)=e^{itH}\mathbf{J}e^{-itH}$. At high temperatures, the Hamiltonian and the current may be given by the sums of individual-particle contributions:

$$H=\sum_i H_i,\qquad \mathbf{J}=\sum_i \mathbf{j}_i. \tag{6.13}$$

Then, the conductivity becomes

$$\sigma=\beta\frac{1}{V}\int_0^\infty\left\langle\sum_i \mathbf{j}_i\mathbf{j}_i(t)+\sum_{i\neq i'}\mathbf{j}_i\mathbf{j}_{i'}(t)\right\rangle dt. \tag{6.14}$$

Here, we have separated the product of the current vectors into two parts.

The statistical average of Eq. (6.14) can be taken using a single-particle state $|k\rangle$ and the corresponding energy $\varepsilon_k{}^0$. As we have discussed in Section 4.3, the grand ensemble operation $\sum_N \prod_{\mathbf{k}}$ can be replaced by $\prod_{\mathbf{k}} \sum_{N_k}$, where N_k is the occupation number. For electrons, N_k is 0 or 1. Furthermore, we recall that $z \exp -\beta\varepsilon_k{}^0 = (1 + z \exp -\beta\varepsilon_k{}^0) f^{(0)}(k)$, where $f^{(0)}(k)$ is the Fermi distribution function.

The first integrand in Eq. (6.14) can be written as follows:

$$\frac{1}{\Xi} \prod_{\mathbf{k}} \sum_{N_k} z^{N_k} (\exp -\beta\varepsilon_k{}^0) N_k \langle k | \mathbf{j}_i \mathbf{j}_i(t) | k \rangle$$

$$= \frac{1}{\Xi} \prod_{\mathbf{k}} (1 + z \exp -\beta\varepsilon_k{}^0) f^{(0)}(k) \langle k | \mathbf{j}_i \mathbf{j}_i(t) | k \rangle,$$

where

$$\Xi = \prod_{\mathbf{k}} (1 + z \exp -\beta\varepsilon_k{}^0)$$

is the grand partition function. Thus, the first term can be expressed as

$$\mathrm{tr}[\mathbf{jj}(t) f^{(0)}(H)], \tag{6.15}$$

where tr means the trace over a complete set of one-electron states and

$$f^{(0)}(H) = \frac{ze^{-\beta H}}{1 + ze^{-\beta H}} \tag{6.16}$$

is the Fermi occupation-number operator.

Similarly, the second integrand is

$$\frac{1}{\Xi} \prod_{\mathbf{k},\mathbf{q}} \sum_{N_k, N_q} z^{N_k} (\exp -\beta\varepsilon_k{}^0) z^{N_q} (\exp -\beta\varepsilon_q{}^0) N_k N_q$$

$$\times [\langle k | \mathbf{j} \; k \rangle \langle q | \mathbf{j} | q \rangle - \langle k | \mathbf{j} | q \rangle \langle q | \mathbf{j}(t) | k \rangle].$$

Here, $\mathbf{k}$ is not supposed to be equal to $\mathbf{q}$, but the term $\mathbf{k} = \mathbf{q}$ may be included because of cancellation. Thus, the above expression becomes

$$-\mathrm{tr}[\mathbf{j} f^{(0)}(H) \mathbf{j}(t) f^{(0)}(H)]. \tag{6.17}$$

Combining the two contributions (6.15) and (6.17), we arrive at

$$\sigma = \beta \frac{1}{V} \int_0^\infty dt \, \mathrm{tr}\left[\mathbf{j} \frac{z^{-1} e^{\beta H}}{1 + z^{-1} e^{\beta H}} \mathbf{j}(t) \frac{ze^{-\beta H}}{1 + ze^{-\beta H}}\right]. \tag{6.18}$$

One can shift $e^{\beta H}$ to the right of $\mathbf{j}(t)$ any number of times by using each time a new variable such as $\tau = t - (i\beta H/\omega)$ and shifting the path of integration.

Thus, we obtain the result due to Chester and Thellung (*19*):

$$\sigma = -\frac{1}{V} \mathrm{tr}\left\{\frac{\partial f^{(0)}}{\partial H} \int_0^\infty dt \, \frac{1}{2} [\mathbf{j}(t)\mathbf{j} + \mathbf{j}\mathbf{j}(t)]\right\}. \tag{6.19}$$

Chester and Thellung further reduced this conductivity formula to a more familiar form for the case when the electrons are scattered elastically by isolated impurities or phonons. In such a case, the total Hamiltonian can be expressed as follows:

$$H = H_0 + \lambda H_1. \tag{6.20}$$

Here, λ is a dimensionless parameter indicating the strength of the scattering potential. Their result, which is correct to order λ^2, is

$$\sigma = -\frac{1}{V} \sum_{\mathbf{k}} \frac{\partial f^{(0)}}{\partial \varepsilon_k{}^0} \langle k | \mathbf{j} | k \rangle \langle k | \mathbf{j} | k \rangle \tau(\varepsilon_k{}^0), \tag{6.21}$$

where $\tau(\varepsilon)$, the relaxation time, includes λ. Replacing the $\mathbf{k}$-summation by integration, we find that this expression agrees with Eq. (5.7) of Chapter 12.

The derivation of Eq. (6.21) is not very simple. On the other hand, this formula can be derived easily by a classical method. Thus it might look as if the correlation function method is not very useful. This view is not correct, however, because the present method shows clearly the conditions under which the formula is derived. Moreover, we can now aim at obtaining higher-order formulas.

13.7.* Cluster Expansion of Time Correlation Functions

If we use the current vector of Eq. (5.4) in Eq. (6.1), the electrical conductivity is determined by the two types of momentum correlation:

$$\begin{aligned} \psi_{11} &= N \langle \mathbf{p}_1(0) \mathbf{p}_1(t) \rangle, \\ \psi_{12} &= N(N-1) \langle \mathbf{p}_1(0) \mathbf{p}_2(t) \rangle, \end{aligned} \tag{7.1}$$

where ψ_{11} represents the self-correlation and ψ_{12} the correlation between a pair of particles. Montroll and Ward (*23*) showed that these correlation functions may be expressed in terms of the distribution functions of imaginary arguments. In other words, the statistical averages entering Eq. (7.1) can be carried out in terms of special distribution functions.

We note that

$$\psi_{11}(t) = \frac{N}{Z_N} \mathrm{tr}[e^{-\beta H_0} e^{-itH_0/\hbar} \mathbf{p}_1(0) e^{itH_0/\hbar} \mathbf{p}_1(0)], \tag{7.2}$$

where

$$Z_N = \mathrm{tr}\, e^{-\beta H_0} \tag{7.3}$$

is the partition function. Thus, using the propagator

$$K(\mathbf{p}'', s_2 ; \mathbf{p}', s_1) = \sum_j e^{-(s_2 - s_1)\varepsilon_j} \psi_j^*(\mathbf{p}') \psi_j(\mathbf{p}''), \tag{7.4}$$

we can write

$$\psi_{11}(t) = \frac{N}{Z_N}\iint \mathbf{p}_1'\mathbf{p}_1'' K\left(\beta, \mathbf{p}''; \beta + \frac{it}{\hbar}, \mathbf{p}'\right) \times K\left(\beta + \frac{it}{\hbar}, \mathbf{p}'; 0, \mathbf{p}''\right) d\mathbf{p}_1'^N \mathbf{p}_1''^N, \tag{7.5}$$

where $\mathbf{p}'$ and $\mathbf{p}''$ of the arguments of K represent all the momenta of N particles over which the integrations are carried out.

It is then appropriate to introduce a singlet distribution function

$$\rho_1^{(N)}(\mathbf{p}_1', \mathbf{p}_1''; t) = \frac{N}{Z_N}\int\cdots\int K\left(\beta, \mathbf{p}''; \beta + \frac{it}{\hbar}, \mathbf{p}'\right) \times K\left(\beta + \frac{it}{\hbar}, \mathbf{p}'; 0, \mathbf{p}''\right) d\mathbf{p}_2'\cdots d\mathbf{p}_N'\, d\mathbf{p}_2''\cdots d\mathbf{p}_N'' \tag{7.6}$$

such that

$$\psi_{11}(t) = \iint \mathbf{p}_1'\mathbf{p}_1''\rho_1^{(N)}(\mathbf{p}_1', \mathbf{p}_1''; t)\, d\mathbf{p}_1'\, d\mathbf{p}_1''. \tag{7.7}$$

The singlet distribution function may be redefined in the grand ensemble, and may be expanded in a cluster series. Since we have already discussed these in Chapters 6 and 11, we shall omit the details. The cluster expansion is

$$\rho_1(\mathbf{p}_1', \mathbf{p}_1''; t) = \Xi^{-1}\sum_N \rho_1^{(N)} Z_N z^N = \sum_{l=1}^{\infty} b_l^{(11)}(\mathbf{p}_1', \mathbf{p}_1''; t) z^l, \tag{7.8}$$

where the b's are cluster expansion coefficients, Ξ is the grand partition function, and z is the absolute activity. We obtain

$$\psi_{11}(t) = \iint \mathbf{p}_1'\mathbf{p}_1'' \sum_{l=1}^{\infty} b_l^{(11)}(\mathbf{p}_1', \mathbf{p}_1''; t) z^l\, d\mathbf{p}_1'\, d\mathbf{p}_1''. \tag{7.9}$$

Similarly, we can prove

$$\psi_{12}(t) = \iint \mathbf{p}_1'\mathbf{p}_2'' \sum_{l=1}^{\infty} b_l^{(12)}(\mathbf{p}_1', \mathbf{p}_2''; t) z^l\, d\mathbf{p}_1'\, d\mathbf{p}_2''. \tag{7.10}$$

13.8.* Master Equation

To discuss irreversibility in quantum mechanical systems, Pauli derived in 1928 the so-called master equation. As we have seen in Chapter 2, the master equation has an elegant structure with irreversible character. Therefore, its derivation from reversible dynamics is very important in statistical mechanics.

Pauli's derivation rests upon the random-phase assumption on the eigenfunctions for all times. This assumption is analogous to the molecular chaos hypothesis for the classical Boltzmann equation, and its repeated use is not satisfactory. In 1955, Van Hove initiated a new era for quantum statistical irreversibility theory by removing the random-phase assumption and showing that a formal perturbation is important in the time evolution of a system starting with certain specified initial states.

Since that time, various improvements have been made by Van Hove (*7*) himself and others, as mentioned at the beginning of this chapter (*8–10*). In what follows, we shall give only a rough, but plausible derivation of the master equation (see the General References).

Let us write down the iterative solution of Eq. (1.17) to second order. We obtain

$$\rho = \rho_0 - \frac{i}{\hbar}\int_{-\infty}^{t}\left(\exp\frac{-i(t-t')H_0}{\hbar}\right)[H_1, \rho_0]\exp\frac{i(t-t')H_0}{\hbar}\,dt' - \left(\frac{1}{\hbar}\right)^2\int_{-\infty}^{t}dt'\int_{-\infty}^{t'}\left(\exp\frac{-i(t-t')H_0}{\hbar}\right)\left[H_1, \left(\exp\frac{-i(t'-t'')H_0}{\hbar}\right)\right. \times [H_1, \rho(t'')]\left.\exp\frac{i(t'-t'')H_0}{\hbar}\right]\exp\frac{i(t-t')H_0}{\hbar}\,dt''. \tag{8.1}$$

Let us assume that ρ_0 is diagonal. Then, the off-diagonal elements of $[H_1, \rho_0]$ are zero. By differentiation, we find for the diagonal element $(d\rho/dt)_{jj}$

$$\left(\frac{d\rho}{dt}\right)_{jj} = -\frac{1}{\hbar^2}\int_{-\infty}^{t}dt''\left[H_1, \left(\exp\frac{-i(t-t'')H_0}{\hbar}\right) \times [H_1, \rho(t'')]\exp\frac{i(t-t'')H_0}{\hbar}\right]_{jj}, \tag{8.2}$$

where we have taken the diagonal elements jj. This equation is exact. Since $[H_1, \rho_0]_{jj} = 0$, the off-diagonal terms of $d\rho/dt$ must be of order H_1. Because the right-hand side of Eq. (8.2) is already of order ${H_1}^2$, we approximate $\rho(t'')$ by the diagonal matrix ρ:

$$\rho_{jk} = \rho_{jj}\,\delta_{jk} + O(H_1). \tag{8.3}$$

We obtain

$$\frac{d\rho_{jj}}{dt} = \frac{2}{\hbar^2}\int_{-\infty}^{t}\sum_k{}' |H_{1jk}|^2(\cos\omega_{jk}\tau)[\rho_{kk}(t-\tau) - \rho_{jj}(t-\tau)]\,d\tau, + \text{off-diagonal contribution.} \tag{8.4}$$

where the summation omits the case $k = j$ and where

$$\omega_{jk} = \frac{E_j - E_k}{\hbar}, \tag{8.5}$$

E_j being the jth energy level of H_0.

Because we are keeping our calculations to order $H_1{}^2$, we expand the elements $\rho_{kk}(t-\tau)$ and $\rho_{jj}(t-\tau)$ in a Taylor series

$$\rho_{jj}(t-\tau) = \rho_{jj}(t) - \tau\frac{d\rho_{jj}}{dt} + \cdots. \tag{8.6}$$

Since $d\rho_{jj}/dt$ is already of order $H_1{}^2$, we may neglect the second term $(-\tau\, d\rho_{ij}/dt)$ and higher terms in the integrand of Eq. (8.4). This is possible as long as the time interval τ is sufficiently small so that

$$\frac{\tau^2 H_1{}^2}{\hbar^2} \ll 1. \tag{8.7}$$

Thus, under this condition, we find

$$\frac{d\rho_{jj}}{dt} = \frac{2}{\hbar^2}\sum_k{}' |H_{1_{jk}}|^2 \frac{\sin(\omega_{jk}t)}{\omega_{jk}}[\rho_{kk}(t) - \rho_{jj}(t)]. \tag{8.8}$$

Let us now take the thermodynamic limit

$$N \to \infty, \qquad V \to \infty, \qquad n = N/V = \text{finite}. \tag{8.9}$$

We also let $t \to \infty$ and $H_1 \to 0$ such that $tH_1{}^2$ remains finite. Then,

$$\frac{\sin(\omega_{jk}t)}{\omega_{jk}} \to \pi\hbar\,\delta(\hbar\omega_{jk}). \tag{8.10}$$

We arrive at

$$\frac{d\rho_{jj}}{dt} = \frac{2\pi}{\hbar}\sum_k{}' |H_{1_{jk}}|^2\,\delta(\varepsilon_j - \varepsilon_k)[\rho_{kk} - \rho_{jj}], \tag{8.11}$$

where $(2\pi/\hbar)|H_{1_{jk}}|^2\,\delta(\varepsilon_j - \varepsilon_k)$ may be replaced by a transition probability a_{jk}. Thus, we have arrived at the master equation under fairly general conditions.

As we have discussed in Section 2.7, the master equation has irreversible character. The above derivation might give the wrong impression that the general perturbation solution of the Bloch equation introduced irreversibility. It is left to the reader to investigate this point.

Due to multiple time integrations, the contributions from off-diagonal elements of the density matrix are in general not negligible. Certain postulates, such as Van Hove's diagonal singularity condition, must be made on the perturbation Hamiltonian in order to obtain the master equation.

Problems

1. For the time correlation function defined by

$$\langle X(t)X(t+s)\rangle = \lim_{\tau\to\infty}\frac{1}{\tau}\int_{-\tau/2}^{\tau/2} X(t)X(t+s)\,dt,$$

show that

$$\langle X(t)X(t+s)\rangle = \frac{1}{2\pi}\int_0^\infty D(\omega)\cos\omega s\,d\omega,$$

where

$$D(\omega) = \lim_{\tau\to\infty}\frac{2}{\tau}\left|\int_{-\tau/2}^{\tau/2} X(t)e^{-i\omega t}\,dt\right|^2.$$

(Wiener–Khinchin theorem). Assume X to vanish for sufficiently large arguments.

2. From the above relation, prove

$$\int_0^\infty \langle X(t)X(t+s)\rangle\,ds = \tfrac{1}{4}D(0+)$$

3. Taking the Fourier transform of the Langevin equation

$$\frac{dv}{dt} + \gamma v = X(t),$$

and using the Wiener–Khinchin theorem (Problem 1), derive the relation

$$\langle v^2(t)\rangle = \frac{1}{2\pi}\int_0^\infty \frac{D(\omega)}{\gamma^2+\omega^2}\,d\omega.$$

4. In the above problem, show that, if the spectrum is white,

$$D(\omega) = \text{const},$$

and if $\langle v^2\rangle = kT/m$, then

$$\gamma = \frac{m}{kT}\int_0^\infty \langle X(t)X(t+s)\rangle\,ds.$$

5. Using the solution of the Langevin equation with the condition $\mathbf{v} = \mathbf{v}_0$ at $t = 0$, show that

$$\int_0^\infty ds\langle F(t)F(t+s)\rangle = \gamma e^{-2\gamma t}\left(v_0{}^2 - \frac{kT}{m}\right),$$

where F is the total force per unit mass (that is, acceleration), given by

$$F(t) = X(t) - \gamma v(t).$$

From this result, observe that in the limit $t \to \infty$ the time correlation of the total force vanishes.

6. Show that (*24*).

$$\langle X(t)v(t+s)\rangle = \frac{1}{4\pi}\int_{-\infty}^{\infty} \frac{D(\omega)}{\gamma - i\omega} e^{i\omega t}\, d\omega;$$

$$\langle X(t+s)v(t)\rangle = \frac{1}{4\pi}\int_{-\infty}^{\infty} \frac{D(\omega)}{\gamma - i\omega} e^{-i\omega t}\, d\omega;$$

$$\langle v(t)v(t+s)\rangle = \frac{1}{4\pi}\int_{-\infty}^{\infty} \frac{D(\omega)}{\gamma^2 + \omega^2} e^{i\omega t}\, d\omega.$$

7. Apply the above results to a white spectrum and prove

$$\langle F(t)F(t+s)\rangle = \langle X(t)X(t+s)\rangle - \frac{\gamma kT}{m} e^{-\gamma t}$$

and confirm the conclusion of Problem 5.

8. Using a narrowband approximation, Eq. (6.11) of Chapter 7 may be reduced to

$$\langle X^2\rangle = \frac{D(\omega)}{2\pi}\Delta\omega.$$

Using this expression and Eqs. (6.1) and (6.11), derive

$$\langle J^2\rangle = 2\frac{V}{\pi} kT\sigma(\omega)\,\Delta\omega.$$

This relation is known as the *Nyquist relation.*

GENERAL REFERENCES

W. E. Brittin, B. W. Downs, and J. Downs (eds.), "Lectures in Theoretical Physics," Vol. 3. Wiley (Interscience), New York, 1961.

E. G. D. Cohen (ed.), "Fundamental Problems in Statistical Mechanics," North-Holland, Amsterdam, 1962.

REFERENCES

1. H. S. Green, "The Molecular Theory of Fluids." North-Holland, Amsterdam, 1952.
2. J. G. Kirkwood, *J. Chem. Phys.* **14**, 180 (1946).
3. N. N. Bogoliubov, *J. Phys. (USSR)* **10**, 265 (1946).

4. G. Sandri, *Ann. Physics* **24**, 332 (1963).
5. H. B. Callen and T. A. Welton, *Phys. Rev.* **83**, 34 (1951).
6. H. B. Callen and R. F. Greene, *Phys. Rev.* **86**, 702 (1952); **88**, 1387 (1952).
7. L. Van Hove, *Physica* **21**, 517 (1955); **23**, 441 (1957); **25**, 268 (1959).
8. E. Verboven, *Physica* **26**, 1091 (1960).
9. A. Janner, *Helv. Phys. Acta* **35**, 47 (1962).
10. R. Brout and I. Prigogine, *Physica* **22**, 621 (1956).
11. R. Balescu and I. Prigogine, *Physica* **25**, 281, 302 (1959).
12. I. Prigogine and P. Resibois, *Physica* **24**, 795 (1958); **27**, 629 (1961).
13. P. Resibois, *Physica* **27**, 541 (1961).
14. H. Nakano, *Progr. Theoret. Phys.* **15**, 77 (1956).
15. H. Mori, *J. Phys. Soc. Japan* **11**, 1029 (1956).
16. H. Mori, *Phys. Rev.* **111**, 694 (1958).
17. R. Kubo, *J. Phys. Soc. Japan* **12**, 570 (1957).
18. W. Kohn and J. M. Luttinger, *Phys. Rev.* **108**, 590 (1957).
18a. M. Lax, *Phys. Rev.* **109**, 1921 (1958).
19. G. V. Chester and A. Thellung, *Proc. Phys. Soc.* **73**, 745 (1959).
20. J. S. Langer, *Phys. Rev.* **120**, 714 (1960); **124**, 1003 (1961); **127**, 5 (1962); **128**, 110 (1962).
21. E. Verboven, *Physica* **26**, 1091 (1960).
22. E. W. Montroll, *Rend. Scuola Inst. Fis.* "*Enrico Fermi*," Varenna (1960).
23. E. W. Montroll and J. C. Ward, *Physica* **25**, 423 (1959).
24. A. Suddaby and P. Gray, *Proc. Phys. Soc.* **75**, 109 (1960).

14

Second Quantization

In 1927, Dirac explained the dual, particle and wave, character of light by showing how photons could arise from electromagnetic fields through a quantization process. Therefore, this process, known as field quantization and as second quantization, started with particles obeying Bose statistics. One year later, Wigner and Jordan extended the method to Fermi statistics.

In many-body theories, the concept of excitons plays an important role. In this respect, the method of second quantization is a powerful tool, as Landau and Bogoliubov showed in their theories of liquid helium.

In this chapter, we shall outline how the method of second quantization is applied to statistical many-body problems. In the next chapter, we shall use this method in combination with that of the Green's functions.

In this and the next chapter, we shall choose the units such that

$$\hbar = 1, \qquad 2m = 1,$$

where m is the particle mass. At first sight, it might look as though such a choice could cause confusion. However, this will not be so as long as we remember the units. Moreover, we will find that theoretical expressions are greatly simplified with this choice and that they can be easily written down and remembered.

In terms of these units, some typical physical quantities appear as shown in Table 14.1.

Sometimes, the unit of the electronic charge e is chosen such that $e^2 = 2$. When this is done, we measure the energy in rydbergs and the length by the Bohr radius. The energy $e^2/2a_0$ is the ionization energy of the hydrogen atom.

TABLE 14.1.

Physical quantity	Old units	New units	Dimension
Momentum	$\hbar k$	k	L^{-1}
Energy	$\hbar^2k^2/2m$	k^2	L^{-2}
Bohr radius a_0	$\hbar^2/me^2$	$2/e^2$	L
Rydberg energy (13.61 eV)	$e^2/2a_0$	$e^4/4$	L^{-2}

14.1. Number Operator

A system may be described in terms of the occupation numbers $n_{\mathbf{k}}$ of the states $\mathbf{k}$ ($n_{\mathbf{k}} = 0, 1, \ldots$). The state vector will then be a function of these numbers and may be represented by a ket vector

$$|\ldots, n_{\mathbf{k}}, \ldots\rangle.$$

We introduce creation and annihilation operators for bosons by

$$\begin{aligned} a_{\mathbf{k}}{}^{+}|\ldots, n_{\mathbf{k}}, \ldots\rangle &= (n_{\mathbf{k}} + 1)^{1/2}|\ldots, n_{\mathbf{k}} + 1, \ldots\rangle, \\ a_{\mathbf{k}}|\ldots, n_{\mathbf{k}}, \ldots\rangle &= n_{\mathbf{k}}^{1/2}|\ldots, n_{\mathbf{k}} - 1, \ldots\rangle. \end{aligned} \tag{1.1}$$

Here, $a_{\mathbf{k}}{}^{+}$, when represented in matrix form, is the Hermitian conjugate of $a_{\mathbf{k}}$.

In terms of $a_{\mathbf{k}}{}^{+}$ and $a_{\mathbf{k}}$, the number operator $\hat{N}_{\mathbf{k}}$ for the state $\mathbf{k}$ is defined for all particles by

$$\hat{N}_{\mathbf{k}} = a_{\mathbf{k}}{}^{+}a_{\mathbf{k}}. \tag{1.2}$$

This definition comes from the relations

$$\begin{aligned} a_{\mathbf{k}}{}^{+}a_{\mathbf{k}}|\ldots, n_{\mathbf{k}}, \ldots\rangle &= n_{\mathbf{k}}|\ldots, n_{\mathbf{k}}, \ldots\rangle; \\ a_{\mathbf{k}}a_{\mathbf{k}}{}^{+}|\ldots, n_{\mathbf{k}}, \ldots\rangle &= (n_{\mathbf{k}} + 1)|\ldots, n_{\mathbf{k}}, \ldots\rangle. \end{aligned} \tag{1.3}$$

Using Eqs. (1.1), and similar equations for fermions we find the commutation relations between $a_{\mathbf{k}}$ and $a_{\mathbf{k}}{}^{+}$:

$$\begin{aligned} a_{\mathbf{k}}a_{\mathbf{s}}{}^{+} - \theta a_{\mathbf{s}}{}^{+}a_{\mathbf{k}} &= \delta_{\mathbf{ks}}; \\ a_{\mathbf{k}}a_{\mathbf{s}} - \theta a_{\mathbf{s}}a_{\mathbf{k}} &= 0; \\ a_{\mathbf{k}}{}^{+}a_{\mathbf{s}}{}^{+} - \theta a_{\mathbf{s}}{}^{+}a_{\mathbf{k}}{}^{+} &= 0. \end{aligned} \tag{1.4}$$

$\theta = 1$ for bosons and $\theta = -1$ for fermions.

The fermion commutation relations are in conformity with the condition that only one particle can be in a given energy state:

$$\hat{N}_{\mathbf{k}}{}^{2} = \hat{N}_{\mathbf{k}}. \tag{1.5}$$

In terms of the number operators, one can introduce the annihilation and creation operators of a particle of arbitrary momentum at a position $\mathbf{r}$. Since the probability amplitude of a particle with momentum $\mathbf{k}$ is

$$\frac{1}{V^{1/2}} \exp(i\mathbf{k}\cdot\mathbf{r}),$$

we have

$$\begin{aligned} \psi_s(\mathbf{r}) &= \frac{1}{V^{1/2}} \sum_{\mathbf{k},\sigma} a_{\mathbf{k}} \exp(i\mathbf{k}\cdot\mathbf{r})\, \delta_{\sigma,s}, \\ \psi_s^{+}(\mathbf{r}) &= \frac{1}{V^{1/2}} \sum_{\mathbf{k},\sigma} a_{\mathbf{k}}^{+} \exp(-i\mathbf{k}\cdot\mathbf{r})\, \delta_{\sigma,s}, \end{aligned} \tag{1.6}$$

where s represents the spin coordinates, V is the volume, and δ is the Kronecker delta. The two field operators ψ_s and ψ_s^{+} satisfy the commutation relations analogous to Eqs. (1.4) . In the bracket notation, the relations are

$$\begin{aligned} &[\psi_s(\mathbf{r}), \psi_t^{+}(\mathbf{r}')]_{\mp} = \delta(\mathbf{r}-\mathbf{r}')\delta_{s,t}, \\ &[\psi_s(\mathbf{r}), \psi_t(\mathbf{r}')]_{\mp} \quad = [\psi_s^{+}(\mathbf{r}), \psi_t^{+}(\mathbf{r}')]_{\mp} = 0, \end{aligned} \tag{1.7}$$

where $\delta(\mathbf{r}-\mathbf{r}')$ is the Dirac delta-function.

The number density of a particle with spin coordinate s at a position $\mathbf{r}$ is given by

$$n_s(\mathbf{r}) = \psi_s^{+}(\mathbf{r})\psi_s(\mathbf{r}). \tag{1.8}$$

The total number of particles in the spin state s is

$$\begin{aligned} \hat{N}_s &= \int \psi_s^{+}(\mathbf{r})\psi_s(\mathbf{r})\, d\mathbf{r} \\ &= \sum_{\mathbf{k}} a_{\mathbf{k}s}^{+} a_{\mathbf{k}s}. \end{aligned} \tag{1.9}$$

The thermal average of the occupation number in the grand ensemble is

$$\langle \hat{N}_{\mathbf{k}} \rangle = \langle a_{\mathbf{k}}^{+} a_{\mathbf{k}} \rangle = \frac{\operatorname{tr}[e^{-\beta\mathscr{H}} a_{\mathbf{k}}^{+} a_{\mathbf{k}}]}{\operatorname{tr} e^{-\beta\mathscr{H}}}, \tag{1.10}$$

where

$$\mathscr{H} = H - \mu N. \tag{1.11}$$

We recall that the trace, if it exists, is invariant under cyclic permutation of operators. Thus,

$$\operatorname{tr}[e^{-\beta\mathscr{H}} a_{\mathbf{k}}^{+} a_{\mathbf{k}}] = \operatorname{tr}[a_{\mathbf{k}} e^{-\beta\mathscr{H}} a_{\mathbf{k}}^{+}] = \operatorname{tr}[e^{-\beta\mathscr{H}} a_{\mathbf{k}} e^{-\beta\mathscr{H}} a_{\mathbf{k}}^{+} e^{\beta\mathscr{H}}]. \tag{1.12}$$

If ψ is an eigenfunction for a system of noninteracting particles, we have [see Eq. (3.11) of Chapter 15]

$$(\exp -\beta\mathscr{H}_0) a_{\mathbf{k}}^{+} (\exp \beta\mathscr{H}_0)\psi \cdot [\exp -\beta(\varepsilon_k^{0} - \mu)] a_{\mathbf{k}}^{+}\psi. \tag{1.13}$$

Therefore, the distribution function $f^{(0)}(k)$ is given by

$$\begin{aligned} \operatorname{tr}[\exp -\beta\mathscr{H}_0) a_{\mathbf{k}}^{+} a_{\mathbf{k}}] &= [\exp -\beta(\varepsilon_{\mathbf{k}}^{0} - \mu)] \operatorname{tr}[(\exp -\beta\mathscr{H}_0) a_{\mathbf{k}} a_{\mathbf{k}}^{+}] \\ &= [\exp -\beta(\varepsilon_{\mathbf{k}}^{0} - \mu)] \operatorname{tr}[(\exp -\beta\mathscr{H}_0)(1 + \theta a_{\mathbf{k}}^{+} a_{\mathbf{k}})], \end{aligned}$$

in other words,

$$f^{(0)}(k) = [\exp -\beta(\varepsilon_k^{0} - \mu)][1 + \theta f^{(0)}(k)]. \tag{1.14}$$

We have arrived at

$$f^{(0)}(k) = \frac{1}{[\exp \beta(\varepsilon_k^{0} - \mu)] - \theta}. \tag{1.15}$$

14.2. Interaction Hamiltonian

Using the number density operator, we can write the kinetic energy as follows:

$$H_0 = \sum_{\mathbf{k}} \int \psi_{\mathbf{k}}^{+}(\mathbf{r})(-\nabla^2)\psi_{\mathbf{k}}(\mathbf{r})\, d\mathbf{r} = \sum_{\mathbf{k}} k^2 a_{\mathbf{k}}^{+} a_{\mathbf{k}}. \tag{2.1}$$

If the interaction is spin-independent, the interaction Hamiltonian is

$$H_1 = \frac{1}{2} \int \sum_{s,s'} \phi(|\mathbf{r}-\mathbf{r}'|) n_s(\mathbf{r}) n_{s'}(\mathbf{r}')\, d\mathbf{r}\, d\mathbf{r}'. \tag{2.2}$$

The right-hand side can be rewritten in terms of the field operators $\psi_s(\mathbf{r})$ and $\psi_s^{+}(\mathbf{r})$ as follows:

$$\begin{aligned} H_1 &= \frac{1}{2} \int \sum_{s,s'} [\psi_s^{+}(\mathbf{r})\phi(|\mathbf{r}-\mathbf{r}'|)\, \delta_{s,s'} \psi_{s'}(\mathbf{r}') \\ &\quad + \psi_s^{+}(\mathbf{r})\psi_{s'}^{+}(\mathbf{r}')\phi(|\mathbf{r}-\mathbf{r}'|)\psi_{s'}(\mathbf{r}')\psi_s(\mathbf{r})]\, d\mathbf{r}\, d\mathbf{r}' \\ &= \frac{1}{2}\phi(0)N + \frac{1}{2} \sum_{s,s'} \int \psi_s^{+}(\mathbf{r})\psi_{s'}^{+}(\mathbf{r}')\phi(|\mathbf{r}-\mathbf{r}'|)\psi_{s'}(\mathbf{r}')\, \psi_s(\mathbf{r})\, d\mathbf{r}\, d\mathbf{r}'. \end{aligned} \tag{2.3}$$

The first term is a constant and can be omitted from our consideration of H_1. Then, the interaction Hamiltonian becomes

$$\begin{aligned} H_1 &= \frac{1}{2} \sum_{s,s'} \int [\psi_s^{+}(\mathbf{r})\psi_{s'}^{+}(\mathbf{r}')\phi(|\mathbf{r}-\mathbf{r}'|)\psi_{s'}(\mathbf{r}')\psi_s(\mathbf{r})]\, d\mathbf{r}\, d\mathbf{r}' \\ &= \frac{1}{2V} \sum_{\mathbf{k},\sigma} (\mathbf{k}_1\mathbf{k}_2|u|\mathbf{k}_3\mathbf{k}_4) a_{\mathbf{k}_1\sigma}^{+} a_{\mathbf{k}_2\sigma'}^{+} a_{\mathbf{k}_4\sigma'} a_{\mathbf{k}_3\sigma}. \end{aligned} \tag{2.4}$$

Here,

$$(\mathbf{k}_1\mathbf{k}_2|u|\mathbf{k}_3\mathbf{k}_4) = \frac{1}{V}\int [\exp(-i\mathbf{k}_1\cdot\mathbf{r} - i\mathbf{k}_2\cdot\mathbf{r}')]$$

$$\times \phi(|\mathbf{r}-\mathbf{r}'|)\exp(i\mathbf{k}_3\cdot\mathbf{r} + i\mathbf{k}_4\cdot\mathbf{r}')\,d\mathbf{r}\,d\mathbf{r}'. \qquad (2.5)$$

The interaction Hamiltonian can be further simplified when the potential ϕ depends only on distance, as we have already indicated in the above expressions. Let us introduce center-of-gravity coordinates and relative coordinates as follows:

$$\mathbf{r}_1 = \tfrac{1}{2}(\mathbf{r}+\mathbf{r}'); \qquad \mathbf{r}_2 = \mathbf{r}-\mathbf{r}'.$$

When the integration on the right-hand side of Eq. (2.5) is performed with respect to $\mathbf{r}$, $\mathbf{r}'$, we can find the interaction Hamiltonian in the following form:

$$H_1 = \frac{1}{2V}\sum_{\substack{\mathbf{k}\sigma\mathbf{k}'\\ \sigma'\mathbf{q}}} u(q)a^+_{\mathbf{k}+\mathbf{q},\sigma}a^+_{\mathbf{k}'-\mathbf{q},\sigma'}a_{\mathbf{k}',\sigma'}a_{\mathbf{k},\sigma}, \qquad (2.6)$$

where

$$u(q) = \int \phi(\mathbf{r})\,[\exp -i\mathbf{q}\cdot\mathbf{r}]\,d\mathbf{r}. \qquad (2.7)$$

The right-hand side represents the situation where two particles with momentum and spin coordinates $\mathbf{k}$, σ and $\mathbf{k}'$, σ' collide, exchange momentum, $\mathbf{q}$, and leave with momenta $\mathbf{k}+\mathbf{q}$ and $\mathbf{k}'-\mathbf{q}$. In the process, the total momentum is conserved. The summation in Eq. (2.6) is over all $\mathbf{q}$, $\mathbf{k}$, $\mathbf{k}'$, σ, and σ'.

14.3. Lattice Vibrations

The lattice vibrations of solids can be described by a displacement vector $\mathbf{u} = \mathbf{r} - \mathbf{r}_0$ of a volume element from an equilibrium position $\mathbf{r}_0$ to a displaced position $\mathbf{r}$. In a harmonic approximation, the displacement vector satisfies the wave equation

$$\left(\frac{\partial^2}{\partial t^2} - c^2\nabla^2\right)\mathbf{u}(\mathbf{r},t) = 0, \qquad (3.1)$$

where c is the velocity of sound.

This wave equation can be derived from the Lagrangian density

$$\mathscr{L} = \frac{\rho_1}{2}\left[\left(\frac{\partial\mathbf{u}}{\partial t}\right)^2 - c^2(\operatorname{div}\mathbf{u})^2\right], \qquad (3.2)$$

where ρ_1 is the ion density. We shall not try to express the density in the units that are used elsewhere in this chapter, because we are interested in the dependence of the Hamiltonian on the mass of the ions. The momentum $\mathbf{P}$, which is conjugate to $\mathbf{u}$, is given by

$$\mathbf{P} = \rho_1\dot{\mathbf{u}}. \qquad (3.3)$$

In quantum mechanics, $\mathbf{u}$ and $\mathbf{P}$ can be considered as operators. We expand these in Fourier series:

$$\mathbf{u}(\mathbf{r}) = \frac{1}{(\rho_1 V)^{1/2}}\sum_{\mathbf{q}}\mathbf{Q}_{\mathbf{q}}\exp(i\mathbf{q}\cdot\mathbf{r}); \qquad \mathbf{Q}_{\mathbf{q}}{}^* = \mathbf{Q}_{\mathbf{q}},$$

$$\mathbf{P}(\mathbf{r}) = \left(\frac{\rho_1}{V}\right)^{1/2}\sum_{\mathbf{q}}\mathbf{P}_{\mathbf{q}}\exp(-i\mathbf{q}\cdot\mathbf{r}). \qquad (3.4)$$

Then, the total Hamiltonian is found to be

$$H = \int\left[\frac{P^2}{2\rho_1} + \frac{\rho_1 c^2}{2}(\operatorname{div}\mathbf{u})^2\right]d\mathbf{r}$$

$$= \tfrac{1}{2}\sum_{\mathbf{q}}[\mathbf{P}_{\mathbf{q}}\cdot\mathbf{P}_{-\mathbf{q}} + c^2q^2\mathbf{Q}_{\mathbf{q}}\cdot\mathbf{Q}_{-\mathbf{q}}]. \qquad (3.5)$$

We can transform this expression into a diagonal form by using

$$\mathbf{P}_{\mathbf{q}} = \left(\frac{cq}{2}\right)^{1/2} i(\mathbf{b}_{\mathbf{q}}{}^+ - \mathbf{b}_{-\mathbf{q}}),$$

$$\mathbf{Q}_{\mathbf{q}} = \left(\frac{1}{2cq}\right)^{1/2}(\mathbf{b}_{\mathbf{q}} + \mathbf{b}^+_{-\mathbf{q}}). \qquad (3.6)$$

We obtain the harmonic oscillator Hamiltonian:

$$H = \sum_{\mathbf{q}}\omega_q(\mathbf{b}_{\mathbf{q}}{}^+\cdot\mathbf{b}_{\mathbf{q}} + \tfrac{1}{2}), \qquad \omega_q = cq. \qquad (3.7)$$

The displacement vector $\mathbf{u}$ can be expressed in terms of the creation and annihilation operators as follows:

$$\operatorname{div}\mathbf{u} = \frac{i}{(\rho_1 V)^{1/2}}\sum_{\mathbf{q}}(\mathbf{Q}_{\mathbf{q}}\cdot\mathbf{q})\exp(i\mathbf{q}\cdot\mathbf{r})$$

$$= \frac{i}{(\rho_1 V)^{1/2}}\sum_{\mathbf{q}}\left(\frac{1}{2cq}\right)^{1/2}(\mathbf{b}_{\mathbf{q}} + \mathbf{b}^+_{-\mathbf{q}})\cdot\mathbf{q}\exp(i\mathbf{q}\cdot\mathbf{r}). \qquad (3.8)$$

14.4. Phonon Spectrum in a Degenerate Bose Gas

Because of the success of Landau's theory of liquid helium, several important attempts have been developed to derive the energy spectrum from first principles. The first approach was made by Bogoliubov in 1947 (*1*). His theory was based on the existence of a large number of particles in the lowest energy state as a result of his assumption of weak couplings between the boson particles [see also Penrose and Onsager (*2*)].

In 1954, Feynman approached the problem by adopting a symmetric wave function with phonon and roton modes. Later, Feynman and Cohen (*3*) introduced a slightly more satisfactory wave function, and, using a variational method, obtained an energy spectrum which is close to that found by experiments.

In 1957, Lee *et al.* (*4*) developed a theory based on a hard-sphere Bose gas model. In this approach, no assumption on condensates such as that made by Bogoliubov was used. More recently, the author and his collaborators (*5*) approached the problem through the pair distribution function. The Fourier transform of this function can be used to evaluate the energy spectrum of interacting bosons.

In what follows, we discuss Bogoliubov's method. The Hamiltonian is

$$H = \sum_{\mathbf{k}} \varepsilon_k{}^0 a_{\mathbf{k}}{}^+ a_{\mathbf{k}} + \tfrac{1}{2} \sum \phi(|\mathbf{k}_1 - \mathbf{k}_1'|) a_{\mathbf{k}_1}^+ a_{\mathbf{k}_2}^+ a_{\mathbf{k}_2}' a_{\mathbf{k}_1}' \, \delta(\mathbf{k}_1 + \mathbf{k}_2, \mathbf{k}_1' + \mathbf{k}_2'). \tag{4.1}$$

We make use of the Bose–Einstein condensation which takes place at $\mathbf{k} = 0$ if $\phi = 0$. If the interaction is weak and if the temperature is very low, we may still be able to assume that most of the particles are in the state $\mathbf{k} = 0$. In other words, the distribution of the momenta of the atoms is assumed to resemble that of an ideal boson gas. Before we proceed with this Bogoliubov model, we remark that the characteristic phonon spectrum has been obtained by Feynman and others without the above assumption. Moreover, according to a crude estimate by Penrose and Onsager (*2*), only 8% of the total atoms seem to participate in the condensation.

The operator for the total number of particles is given by

$$\hat{N} = \hat{N}_0 + \sum_{\mathbf{k}}{}' a_{\mathbf{k}}{}^+ a_{\mathbf{k}}, \tag{4.2}$$

where $\hat{N}_0$ is the number operator of particles in the state $\mathbf{k} = 0$, and the second term represents the particles in the excited states. We shall treat $\hat{N}_0$ as a c-number and replace it by the total number of particles N. Also, in the interaction Hamiltonian, we shall retain only the terms that involve the particles in the lowest states. The result of these approximations is

$$H = \tfrac{1}{2}N^2\phi_0 + \sum_{\mathbf{k}}{}' (\varepsilon_k{}^0 + N\phi_k) a_{\mathbf{k}}{}^+ a_{\mathbf{k}} + \tfrac{1}{2}N \sum_{\mathbf{k}}{}' \phi_k (a_{\mathbf{k}} a_{-\mathbf{k}} + a_{-\mathbf{k}}^+ a_{\mathbf{k}}{}^+)$$
$$+ \text{higher terms}, \tag{4.3}$$

where the first term comes from the interactions among the particles in the lowest state and is a constant. By neglecting higher terms, the Hamiltonian may be brought into the following form:

$$H - \tfrac{1}{2}N^2\phi_0 = \sum{}' H_k,$$
$$H_k = \alpha_0 (a_{\mathbf{k}}{}^+ a_{\mathbf{k}}) + \tfrac{1}{2}\alpha_1 (a_{\mathbf{k}} a_{-\mathbf{k}} + a_{-\mathbf{k}}^+ a_{\mathbf{k}}{}^+), \tag{4.4}$$

where

$$\alpha_0 = \varepsilon_k{}^0 + N\phi_k, \qquad \alpha_1 = N\phi_k. \tag{4.5}$$

In order to find the energy spectrum, we try to express the Hamiltonian in the form:

$$H = \sum_{\mathbf{k}} \varepsilon_k b_{\mathbf{k}}{}^+ b_{\mathbf{k}} + \text{const.} \tag{4.6}$$

We find that this Hamiltonian satisfies the commutation relations

$$[b_{\mathbf{k}}{}^+, H] = -\varepsilon_k b_{\mathbf{k}}{}^+,$$
$$[b_{\mathbf{k}}, H] = \varepsilon_k b_{\mathbf{k}}, \tag{4.7}$$

for boson operators:

$$[b_{\mathbf{k}}, b_{\mathbf{k}'}^+] = \delta_{\mathbf{k},\mathbf{k}'}. \tag{4.8}$$

Since the original and the diagonalized Hamiltonians are both quadratic, it is clear that the operators $b_{\mathbf{k}}$ and $b_{\mathbf{k}}{}^+$ are related to $a_{\mathbf{k}}$ and $a_{\mathbf{k}}{}^+$ linearly. Thus, we introduce the Bogoliubov transformation

$$b_{\mathbf{k}} = u_{\mathbf{k}} a_{\mathbf{k}} - v_{\mathbf{k}} a_{-\mathbf{k}}^+,$$
$$b_{\mathbf{k}}{}^+ = u_{\mathbf{k}} a_{\mathbf{k}}{}^+ - v_{\mathbf{k}} a_{-\mathbf{k}}. \tag{4.9}$$

From the boson commutation relation, we find

$$[b_{\mathbf{k}}, b_{\mathbf{k}}{}^+] = u_{\mathbf{k}}{}^2 - v_{\mathbf{k}}{}^2 = 1. \tag{4.10}$$

Also, from commutation relations such as

$$[b_{\mathbf{k}}{}^+, H] = u_{\mathbf{k}}(-\alpha_0 a_{\mathbf{k}}{}^+ - \alpha_1 a_{-\mathbf{k}}) - v_{\mathbf{k}}(\alpha_0 a_{-\mathbf{k}} + \alpha_1 a_{\mathbf{k}}{}^+)$$
$$= -\varepsilon_k b_{\mathbf{k}}{}^+$$
$$= -\varepsilon_k (u_{\mathbf{k}} a_{\mathbf{k}}{}^+ - v_{\mathbf{k}} a_{-\mathbf{k}}), \tag{4.11}$$

we arrive at a set of linear equations, the solutions of which exist when the secular equation is satisfied:

$$\begin{vmatrix} \alpha_0 - \varepsilon_k & \alpha_1 \\ \alpha_1 & \alpha_0 + \varepsilon_k \end{vmatrix} = 0. \tag{4.12}$$

Thus, we find the energy eigenvalues in the form

$$\begin{aligned}\varepsilon_k{}^2 &= \alpha_0{}^2 - \alpha_1{}^2 \\ &= (\varepsilon_k{}^0 + N\phi_k)^2 - (N\phi_k)^2. \end{aligned} \tag{4.13}$$

Since $\varepsilon_k{}^0 = k^2/2m$, the eigenvalue for small, positive k may be obtained by neglecting $\varepsilon_k^{0^2}$. Then, it is phononlike:

$$\varepsilon_k = \left(\frac{N\phi_0}{m}\right)^{1/2} k, \tag{4.14}$$

where we have made explicit the dependence on m, the mass, and have assumed that the interaction is repulsive: $\phi_0 > 0$. The sound velocity is given by

$$c = \left(\frac{N\phi_0}{m}\right)^{1/2}.$$

On the other hand, for large values of k, we obtain

$$\varepsilon_k = \varepsilon_k{}^0 + N\phi_k. \tag{4.15}$$

The first term $\varepsilon_k{}^0$ varies as k^2. Therefore, if ϕ_k is small for large k, the energy spectrum is particle-like.

The excitation energy spectrum of liquid helium was shown in Fig. 4.9 (p. 89). The slope of the initial straight line gives the velocity of sound

$$c = 2.39 \times 10^4 \quad \text{cm sec}^{-1}. \tag{4.16}$$

14.5. Electron Gas

As we have discussed in Section 4.6, electrons at 0°K fill all the energy levels up to the Fermi level. In the absence of Coulomb interaction, the system at 0°K would have only the kinetic energy:

$$E_0 = \sum_{\mathbf{k} \le \mathbf{k}_0} k^2 = \tfrac{3}{5} N k_0{}^2, \tag{5.1}$$

where the last equation is obtained by replacing the summation $V^{-1}\sum$ by the integration $2(2\pi)^{-3}\int d\mathbf{k}$, with the factor 2 representing spin degeneracy.

The Coulomb interaction alters this energy. The first correction is the *exchange energy* E_x,

$$E_x = \frac{1}{2V}\sum_{\mathbf{k},\mathbf{q}} \frac{4\pi e^2}{q^2} \langle 0 | a^+_{\mathbf{k}+\mathbf{q}} a^+_{\mathbf{k}'-\mathbf{q}} a_{\mathbf{k}'} a_{\mathbf{k}} | 0 \rangle, \tag{5.2}$$

where $|0\rangle$ is the ground-state wave function,

$$H_0|0\rangle = E_0|0\rangle. \tag{5.3}$$

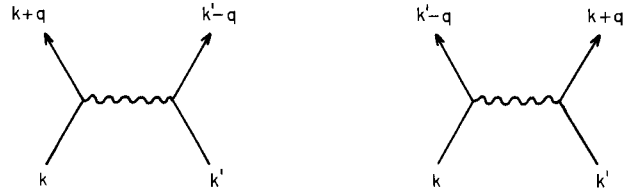


Fig. 14.1. Electron–electron scattering.

As shown in Fig. 14.1, the matrix elements of E_x represent two types of configurations. The first graph is not permissible since $\mathbf{k} + \mathbf{q} = \mathbf{k}$ requires $\mathbf{q} = 0$ which should be excluded. This exclusion is justified by considering a cloud of positive charges in the background. The second graph yields

$$E_x = -\frac{1}{2V}\sum_{\mathbf{q}} \frac{4\pi e^2}{q^2} \sum_{\mathbf{k}} n_{\mathbf{k}+\mathbf{q}} n_{\mathbf{k}}, \tag{5.4}$$

Since

$$n_{\mathbf{k}} = \begin{cases} 1, & k < k_0, \\ 0, & k > k_0, \end{cases} \tag{5.5}$$

we obtain

$$E_x = -\frac{1}{(2\pi)^3} \int_{\substack{|\mathbf{k}+\mathbf{q}|<|\mathbf{k}_0| \\ |\mathbf{k}|<|\mathbf{k}_0|}} \frac{4\pi e^2}{q^2}\, d\mathbf{k}\, d\mathbf{q} \tag{5.6}$$

$$= -\frac{V e^2 k_0{}^4}{4\pi^3}$$

$$= -\frac{3e^2 N k_0}{4\pi}, \tag{5.7}$$

where

$$k_0{}^3 = 3\pi^2 n. \tag{5.8}$$

More convenient expressions are obtained if we use units such that $e^2 = 2$ in addition to $2m = 1$, $\hbar = 1$. Then, the energies per particle are

$$\begin{aligned} \varepsilon_0 &= \frac{E_0}{N} = \frac{3}{5}\left(\frac{9\pi}{4}\right)^{2/3} \frac{1}{r_s{}^2} \approx \frac{2.21}{r_s{}^2}; \\ \varepsilon_x &= \frac{E_x}{N} = \frac{E_x/N}{me^4/2\hbar^2} = -\frac{3}{2\pi}\left(\frac{9\pi}{4}\right)^{1/3} \frac{1}{r_s} = -\frac{0.916}{r_s}, \end{aligned} \tag{5.9}$$

where

$$r_s = \frac{r_0}{a_0}, \qquad \frac{4\pi r_0{}^3}{3} = \frac{1}{n}. \tag{5.10}$$

The radius of the sphere representing the volume per electron is r_0. The dimensionless parameter r_s is the same radius measured in units of the Bohr radius a_0. A method of obtaining Eqs. (5.9) is indicated by Problem 5. The total energy is

$$\varepsilon_g = \frac{2.21}{r_s^2} - \frac{0.916}{r_s} + \varepsilon_c, \tag{5.11}$$

where ε_c is the correlation energy. As we have shown in Eq. (9.11) of Chapter 10, it is approximately given by

$$\varepsilon_c = -0.096 + 0.0622 \log r_s + \cdots. \tag{5.12}$$

The above expression holds for small r_s, that is, for high density.

If the number v of excited particles is small compared to the total number N, the interaction between the excited particles may be neglected. Therefore, the energy levels relative to the ground state will be of the same structure as those of a gas of independent particles. With $k_j > k_0$ and $p_j < k_0$,

$$U = E_0 + \sum_{j=1} [W(k_j) - W(p_j)] + O\left(\frac{v}{N}\right). \tag{5.13}$$

For a free electron, if W is expressed in rydbergs and if $\alpha = (4/9\pi)^{1/2}$,

$$W = \frac{p^2}{\alpha^2 r_s^2}. \tag{5.14}$$

Thus, the specific heat is determined by

$$\left[\left(\frac{dW}{dp}\right)_{p=1}\right]^{-1} = \frac{\alpha^2 r_s^2}{2}, \tag{5.15}$$

where $(4/3)\pi r_s^3$ is the volume per electron in units of the Bohr radius cubed. In ordinary units, the specific heat is

$$\begin{aligned} c_V^0(T, r_s) &= m^{-1}\hbar^2 e^{-4} k^2 T \alpha^2 r_s^2, \\ c_V(T, r_s) &= m^{-1}\hbar^2 e^{-4} k^2 T\{2[(dW/dp)_{p=1}]^{-1}\}. \end{aligned} \tag{5.16}$$

The specific heat is obtained by evaluating dW/dp for high density. The result in Eq. (9.12) of chapter 10 can be expressed as follows:

$$c_V = c_V^0[1 + 0.083 r_s(-\log r_s - 0.203) + \cdots]^{-1}. \tag{5.17}$$

14.6. Electron–Phonon Interaction

Phonons arise in ionic crystals and metals due to the vibrations of ions. Therefore, the interaction between electrons and phonons may be assumed to have the form

$$\frac{1}{r} \times \text{electron charge density} \times \text{ion charge density},$$

where the ion charge density varies with lattice deformation. The volume change induced by a displacement vector $\mathbf{u}$ in the lattice of volume V is V div $\mathbf{u}$, and the change in the ion number density Δn_I due to lattice vibrations is approximately given by

$$\Delta n_I = -n_I \operatorname{div} \mathbf{u}. \tag{6.1}$$

Therefore, for the electron–phonon interactions, we use

$$\begin{aligned} H_I &= \int \frac{(-Ze)en_I}{|\mathbf{r} - \mathbf{r}'|} \psi^+(\mathbf{r})\psi(\mathbf{r})[-\operatorname{div} \mathbf{u}(\mathbf{r}')]\, d\mathbf{r}\, d\mathbf{r}' \\ &= \frac{4\pi e^2 Z n_I}{V} \sum_{\mathbf{q}} \int \frac{\exp i\mathbf{q}\cdot(\mathbf{r} - \mathbf{r}')}{q^2} \psi^+(\mathbf{r})\psi(\mathbf{r})[\operatorname{div} \mathbf{u}(\mathbf{r}')]\, d\mathbf{r}\, d\mathbf{r}'. \end{aligned} \tag{6.2}$$

Here, $-Ze$ is the ion charge. Using Eqs. (3.8) and (1.6) and neglecting the spin, we find

$$H_I = \frac{4\pi Z e^2 n_I}{(\rho_I V)^{1/2}} \sum_{\mathbf{k},\mathbf{q}} (2cq)^{-1/2} \frac{i}{q^2} a^+_{\mathbf{k}+\mathbf{q}} a_{\mathbf{k}} \mathbf{q}\cdot(\mathbf{b}_{\mathbf{q}} + \mathbf{b}^+_{-\mathbf{q}}), \tag{6.3}$$

where $\mathbf{b}$ and $\mathbf{b}^+$ are for phonons and a and a^+ are for electrons.

The right-hand side of Eq. (6.3) consists of the terms that represent the electrons changing their vectors from $\mathbf{k}$ to $\mathbf{k} + \mathbf{q}$ either by absorbing a phonon of momentum $\mathbf{q}$ or by emitting a phonon of momentum $-\mathbf{q}$ (Fig. 14.2).

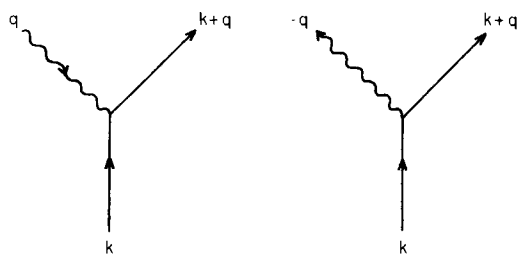


Fig. 14.2. Electron–phonon scattering.

On the right-hand side of Eq. (6.3), each electron is coupled with a cloud of phonons, forming a composite particle. In ionic crystals, longitudinal phonons with optical frequencies produce electric polarization waves which couple strongly with electrons. The resulting composite particles are called *polarons*.

To describe polarons in dielectrics, the polarization vector $\mathbf{P}$ has to be used instead of the displacement vector $\mathbf{u}$. We expand $\mathbf{P}$ in a Fourier series:

$$\mathbf{P} = F\sum_{\mathbf{q}} \varepsilon_{\mathbf{q}}[b_{\mathbf{q}}(\exp i\mathbf{q}\cdot\mathbf{r}) + b_{\mathbf{q}}{}^{+}(\exp -i\mathbf{q}\cdot\mathbf{r})], \tag{6.4}$$

where F is the strength of $\mathbf{P}$ and $\varepsilon_{\mathbf{q}}$ is a unit vector in the direction of $\mathbf{q}$. The electric field $\mathbf{E}$ can be found from $\operatorname{div}\mathbf{D} = \operatorname{div}(\mathbf{E} + 4\pi\mathbf{P}) = 0$. Finally, the interaction Hamiltonian is found as follows:

$$H_1 = -4\pi iFe\sum_{\mathbf{k},\mathbf{q}} q^{-1}(b_{\mathbf{q}}a^{+}_{\mathbf{k}+\mathbf{q}}a_{\mathbf{k}} - b_{\mathbf{q}}{}^{+}a^{+}_{\mathbf{k}-\mathbf{q}}a_{\mathbf{k}}). \tag{6.5}$$

We assume that in the ground state there are no phonons. The ground-state wave function with an electron of momentum $\mathbf{k}$ and no phonon can be represented by $|\mathbf{k}0\rangle$. According to quantum mechanics, the first-order perturbed state is given by

$$|\mathbf{k}, 0\rangle^{(1)} = |\mathbf{k}, 0\rangle + \sum_{\mathbf{q}} |\mathbf{k}-\mathbf{q}; 1\rangle \frac{\langle \mathbf{k}-\mathbf{q}; 1|H_1|\mathbf{k}, 0\rangle}{\varepsilon_k - \varepsilon_{\mathbf{k}-\mathbf{q}} - \omega_{\mathbf{q}}}, \tag{6.6}$$

where $|\mathbf{k}-\mathbf{q}; 1\rangle$ represents the state with one phonon of energy ω_q excited. To second order, the self-energy of a polaron in the weak coupling limit is given by

$$\varepsilon_{\mathbf{k}} = \varepsilon_{\mathbf{k}}{}^{0} - \sum_{\mathbf{q}} \frac{\langle \mathbf{k}-\mathbf{q}; 1|H_1|\mathbf{k}, 0\rangle^2}{(q^2 - 2\mathbf{k}\cdot\mathbf{q} + \omega_{\mathbf{q}})}. \tag{6.7}$$

$\varepsilon_{\mathbf{k}}{}^{0}$ is the kinetic energy.

The sum on the right-hand side of Eq. (6.7) can be evaluated for a slow electron ($\mathbf{k}^2 < \omega_{\mathbf{q}} = q_p{}^2$) by replacing the summation by an integration (*6*, *7*). For longitudinal polarization waves of strength F, Eq. (6.5) gives

$$\langle \mathbf{k}-\mathbf{q}; 1|H_1|\mathbf{k}, 0\rangle^2 = \frac{(4\pi eF)^2}{q^2}. \tag{6.8}$$

Using Eq. (6.7) and replacing summation by integration, we obtain

$$\varepsilon_k = \varepsilon_k{}^{0} - 4e^2F^2\int_{-1}^{1} d(\cos\theta)\int_0^{\infty} dq\,\frac{1}{q^2 - 2\mathbf{k}\cdot\mathbf{q} + q_p{}^2}$$

$$= \varepsilon_k{}^{0} - \alpha\left(\omega_{\mathbf{q}} + \frac{k^2}{12m} + \cdots\right), \tag{6.9}$$

where the electronic mass m has been reintroduced, and where

$$\alpha = 16e^2F^2(2m)^2\int_0^{\infty} dq\,\frac{1}{(q^2 + q_p{}^2)^2} = \frac{16\pi e^2F^2m^2}{q_p{}^3}. \tag{6.10}$$

The energy is given approximately by

$$\varepsilon_k = \frac{(1 - \frac{1}{6}\alpha)k^2}{2m}. \tag{6.11}$$

If $\alpha \ll 6$, the mass of a polaron is given approximately by

$$m^{p}_{\text{eff}} = m(1 + \tfrac{1}{6}\alpha). \tag{6.12}$$

Electron mobility experiments on AgBr show $\alpha \approx 1.60$ and $m_{\text{eff}} = 1.26m$. (*6*, *7*). However, the electron mass m itself should be replaced by an effective mass, which is of the order of $0.20m$. Thus, the effective mass for a polaron is approximately $0.27m$.

The fact that the effective mass of an electron increases due to electron–phonon couplings results in larger electronic specific heats. This conclusion has been confirmed by experiments.

In Eq. (6.2), the electron–ion interaction is a bare Coulomb force. However, in metals, it is necessary to consider a shielding effect by ions. The Fourier transform of the force in such a case can be approximated by

$$u(q) = \frac{4\pi e^2}{q^2 + \mu_0{}^2}, \tag{6.13}$$

where

$$\mu_0 = \left(6\pi^2 n\frac{e^2}{k_0{}^2}\right)^{1/2} \tag{6.14}$$

is the Thomas–Fermi screening constant which appeared in Eq. (6.8) of Chapter 11. For phonons of low wave numbers, we can replace $u(q)$ by $4\pi e^2/\mu_0{}^2$. Then, the interaction Hamiltonian becomes

$$H_1 = i\sum_{\mathbf{k},\mathbf{q}} D(q)a^{+}_{\mathbf{k}+\mathbf{q}}a_{\mathbf{k}}(b_{\mathbf{q}} - b^{+}_{-\mathbf{q}}), \tag{6.15}$$

where the operators $b_{\mathbf{q}}$ and $b_{\mathbf{q}}{}^{+}$ are in the direction of $\mathbf{q}$ (that is, they act on longitudinal phonons), and the constant $D(q)$ is inversely proportional to the square root of the ion density times the velocity of sound.

14.7. Electron–Electron Interaction via Phonons

The Hamiltonian obtained in Eq. (6.15) has a characteristic factor $D(q)$. Let us introduce a canonical transformation and try to find an approximate Hamiltonian in which terms linear in H_1 are eliminated. This transformation is achieved as follows. If

$$H = H_0 + \xi H_1, \tag{7.1}$$

then

$$\tilde{H} = e^{-S}He^{S} = H + [H, S] + \tfrac{1}{2}[[H, S], S] + \cdots. \tag{7.2}$$

We choose S such that

$$\xi H_1 + [H_0, S] = 0. \tag{7.3}$$

Then, the transformed Hamiltonian becomes

$$\tilde{H} = H_0 + \xi[H_1, S] + \cdots, \tag{7.4}$$

where S is given by

$$\langle n|S|m\rangle = \frac{\langle n|H_1|m\rangle}{E_m - E_n}. \tag{7.5}$$

Applying Eq. (7.5) to our problem of finding an electron–electron interaction via phonons (*8*), we note that, at absolute zero, one phonon excitation from the ground state may be considered to be dominant. Thus, if we neglect a dependence of $D(q)$ on q and express S in terms of electron operators, we obtain matrix elements such as

$$\langle 1|S|0\rangle = -iD\sum_{\mathbf{k}} a^{+}_{\mathbf{k}-\mathbf{q}}a_{\mathbf{k}}\frac{1}{\varepsilon_k - \varepsilon_{\mathbf{k}-\mathbf{q}} - \omega_{\mathbf{q}}}. \tag{7.6}$$

Thus, we find an approximate Hamiltonian:

$$\tilde{H} = H_0 + D^2\sum_{\mathbf{q}}\sum_{\mathbf{k}\mathbf{k}'}\frac{\omega_q}{(\varepsilon_k - \varepsilon_{\mathbf{k}-\mathbf{q}})^2 - \omega_{\mathbf{q}}^2}a^{+}_{\mathbf{k}'+\mathbf{q}}a_{\mathbf{k}'}a^{+}_{\mathbf{k}-\mathbf{q}}a_{\mathbf{k}}. \tag{7.7}$$

The second term is of order D^2 and represents an electron–electron interaction via a phonon. (Fig 14.1). Note that this part is negative (attractive) for $|\varepsilon_{\mathbf{k}\pm\mathbf{q}} - \varepsilon_q| < \omega_{\mathbf{q}}$.

Let us assume that in superconductors the electron–electron couplings take place only in a narrow energy range of width $2\omega_{\mathrm{D}}$ about the Fermi surface:

$$\varepsilon_0 - \omega_{\mathrm{D}} < \varepsilon_k, \varepsilon_{\mathbf{k}\pm\mathbf{q}} < \varepsilon_0 + \omega_{\mathrm{D}}, \tag{7.8}$$

where ε_0 is the Fermi energy. The effective Hamiltonian becomes

$$H_{\mathrm{eff}} = -\tfrac{1}{2}\phi\sum_{\mathbf{q}}\sum_{\mathbf{k}\mathbf{k}'}a^{+}_{\mathbf{k}'+\mathbf{q}}a_{\mathbf{k}}a^{+}_{\mathbf{k}-\mathbf{q}}a_{\mathbf{k}}. \tag{7.9}$$

Note that D^2 and ω_{D}, the Debye frequency, are both inversely proportional to ion mass. Consequently, the new potential parameter ϕ, which is positive, is independent of ion mass.

14.8. f-Sum Rule

For a Hamiltonian of the form

$$H = \sum_i p_i^2 + \sum_{i\neq j}\tfrac{1}{2}\phi(|\mathbf{r}_i - \mathbf{r}_j|), \tag{8.1}$$

Placzek (*9*), and later Nozieres and Pines (*10*) showed that

$$\sum_j \frac{1}{k^2}|(n_{\mathbf{k}}^{*})_{j0}|^2\omega_{j0} = N. \tag{8.2}$$

Before giving the proof of this relation, known as the *f-sum rule*, let us introduce

$$S(k, \omega) = \sum_j (n_{\mathbf{k}}^{*})_{j0}^2\,\delta(\omega - \omega_{j0}), \tag{8.3}$$

where $S(k, \omega)$ is called the *dynamic form factor*. Here, $n_{\mathbf{k}}$ is the Fourier transform of the density:

$$n_{\mathbf{k}} = \sum_i \exp(-i\mathbf{k}\cdot\mathbf{r}_i). \tag{8.4}$$

The coefficient $n_{\mathbf{k}}$ is essentially the same as that in Eq. (7.3) of Chapter 9. Note that $n_0 \neq 0$ in the present case. Here, $(n_{\mathbf{k}})_{j0}$ is the matrix element of the density fluctuations between the ground state and the jth excited state, and ω_{j0} is the corresponding excitation frequency.

The f-sum rule (8.2) follows from the relations

$$[n_{\mathbf{k}}, H] = \sum_i (2\mathbf{k}\cdot\mathbf{p}_i - k^2)\exp(-i\mathbf{k}\cdot\mathbf{r}_i), \tag{8.5}$$

and

$$[n_{\mathbf{k}}^{*}, [n_{\mathbf{k}}, H]] = 2Nk^2. \tag{8.6}$$

For the ground state, we have

$$\begin{aligned}[n_{\mathbf{k}}^{*}, [n_{\mathbf{k}}, H]]_{00} &= 2Nk^2 \\ &= \sum\{(n_{\mathbf{k}}^{*})_{0j}[n_{\mathbf{k}}, H]_{j0} - [n_{\mathbf{k}}, H]_{0j}(n_{\mathbf{k}}^{*})_{j0}\}.\end{aligned} \tag{8.7}$$

Using the excitation energy $E_j - E_0 = \omega_{j0}$ and also the relation

$$[n_{\mathbf{k}}, H]_{j0} = \omega_{j0}(k)(n_{\mathbf{k}})_{j0}, \tag{8.8}$$

we can evaluate the double commutator in the following form:

$$\begin{aligned}\sum_j [(n_{\mathbf{k}}^{*})_{0j}(n_{\mathbf{k}})_{j0}\,\omega_{j0}(k) + (n_{\mathbf{k}})_{0j}(n_{\mathbf{k}}^{*})_{j0}\,\omega_{j0}(k)] \\ = 2\sum_j \omega_{j0}(k)|(n_{\mathbf{k}}^{*})_{j0}|^2.\end{aligned} \tag{8.9}$$

This gives Eq. (8.2), or

$$\int_0^\infty d\omega\, S(k,\omega)\omega = Nk^2. \tag{8.10}$$

The correlation function $S(k, \omega)$ characterizes the energy states of a system. For a free particle, we have $(\hbar k)^2/2m$ for the energy of excited states in ordinary units, and

$$S(k,\omega) = N\,\delta(\omega - k^2). \tag{8.11}$$

Correspondingly, the structure factor is expressed by

$$S(k) = 1, \tag{8.12}$$

which means that there is actually no correlation.

On the other hand, if the scattering occurs overwhelmingly at $\omega = \varepsilon(k)$

$$\begin{aligned} S(k,\omega) &= \sum_j (n_{\mathbf{k}}^*)_{j0}^2\, \delta(\omega - \varepsilon(k)) \\ &= NS(k)\,\delta(\omega - \varepsilon(k)), \end{aligned} \tag{8.13}$$

the f-sum rule gives the Feynman expression for the energy spectrum:

$$\varepsilon(k) = \frac{k^2}{S(k)}. \tag{8.14}$$

In these equations, the *static form factor* $S(k)$ is related to $S(k, \omega)$ by

$$S(k) = N^{-1}\int_0^\infty S(k,\omega)\, d\omega. \tag{8.15}$$

Equation (8.14) agrees with Feynman's expression for the excitation spectrum of He II. Note that it has been derived under the assumption that the density fluctuations are the only possible excitations. Note also that Bose statistics has not been used explicitly in this derivation.

14.9. Dielectric Constant of a Plasma

A system of charged particles may be probed by a test charge (*11–13*). The Hamiltonian W for the total system can be written as

$$W = H + H_1, \tag{9.1}$$

where

$$H = \sum_i p_i^2 + \sum_{\mathbf{k}} \frac{2\pi e^2}{k^2}(n_{\mathbf{k}}^* n_{\mathbf{k}} - N), \tag{9.2}$$

and

$$H_1 = e^2 \int \frac{n_T(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\, d\mathbf{r}\, d\mathbf{r}' \tag{9.3}$$

$$= \sum_{\mathbf{k}} \frac{4\pi e^2}{k^2} n_{-\mathbf{k}} I(k) e^{-i\omega t}, \tag{9.4}$$

if the density of the test charge is given by

$$n_T(\mathbf{r}) = \sum_{\mathbf{k}} I(k)[\exp\, i(\mathbf{k}\cdot\mathbf{r} - \omega t)] + \text{c.c.} \tag{9.5}$$

Let us assume that the influence of the test charge is sufficiently small that first-order treatment of the coupling between the test charge and the system is permissible. The wave function of the system is projected onto the eigenstates $|j\rangle$ of the energy E_j of H such that

$$\Psi(t) = \sum a_j(t) e^{-iE_j t}|j\rangle. \tag{9.6}$$

In order to find $\Psi(t)$ by a time-dependent perturbation, we use

$$a_j(-\infty) = \begin{cases} 1, & j = 0 \\ 0, & j \neq 0. \end{cases} \tag{9.7}$$

To first order, we find for $\eta > 0$

$$a_j(t) = \frac{4\pi e^2}{k^2}\left[\frac{I(k)(n_{-\mathbf{k}})_{j0}\, e^{i(-\omega+\omega_{j0})t+\eta t}}{\omega - \omega_{j0} + i\eta} - \frac{I(-k)(n_{\mathbf{k}})_{j0} e^{i(\omega+\omega_{j0})t+\eta t}}{\omega + \omega_{j0} - i\eta}\right], \tag{9.8}$$

where $\omega_{j0} = E_j - E_0$, and η is a parameter which switches on the interactions adiabatically at $t = -\infty$ so that the stationary state is reached at $t = 0$. This parameter is necessary because the interaction, even though it is weak, extends from $t = -\infty$ to $t = \infty$ and the accumulation of its effects could cause complications such as a nonlinear response. The parameter shall be set equal to zero after the calculations.

After we find the wave function, we can evaluate expectation values. The linear response of a plasma to a test charge is expressed by a *dielectric function*, which is obtained from

$$\frac{1}{\varepsilon(k,\omega)} = 1 + \frac{\langle n_{\mathbf{k}}\rangle}{VI(k)} \tag{9.9}$$

(where $\langle n_{\mathbf{k}}\rangle$ is the Fourier transform of the average density fluctuations), because the Maxwell field and source relations are

$$\begin{aligned} &\text{div}\,\mathbf{D} = 4\pi V^{-1} \times \text{test charge}, \\ &\text{div}\,\mathbf{E} = 4\pi V^{-1} \times (\text{test charge} + \text{induced charges}). \end{aligned} \tag{9.10}$$

Taking the Fourier transforms of the longitudinal components, we obtain Eq. (9.9), where V is the volume and $I(k)$ represents the strength of the test charge.

Evaluating $\langle n_{\mathbf{k}} \rangle$, we find

$$\varepsilon(k, \omega)^{-1} = 1 + \sum_j \frac{4\pi e^2}{k^2 V} |(n_{\mathbf{k}})_{j0}|^2 \left(\frac{1}{\omega - \omega_{j0} + i\eta} - \frac{1}{\omega + \omega_{j0} + i\eta} \right). \quad (9.11)$$

We note the mathematical identity

$$\lim_{\eta \to 0} \frac{1}{x \mp i\eta} = \frac{P}{x} \pm i\pi\, \delta(x), \quad (9.12)$$

where P indicates principal value. We obtain

$$\operatorname{Im} \frac{1}{\varepsilon(k, \omega)} = \frac{4\pi^2 e^2}{k^2 V} \sum_j |(n_{\mathbf{k}})_{j0}|^2 [\delta(\omega + \omega_{j0}) - \delta(\omega - \omega_{j0})]. \quad (9.13)$$

If $\omega_{j0} > 0$ and $\omega > 0$, we can use

$$\frac{4\pi^2 e^2}{k^2 V} S(k, \omega) = \operatorname{Im} \frac{1}{\varepsilon(k, \omega)}. \quad (9.14)$$

Then, the f-sum rule for $S(k, \omega)$ gives

$$\int_0^\infty \omega \operatorname{Im} \frac{1}{\varepsilon(k, \omega)}\, d\omega = -4\pi^2 n e^2. \quad (9.15)$$

We can also show that the expectation value of the Coulomb interaction energy in the ground state is given by

$$E_{\text{int}} = -\sum_{\mathbf{k}} \left[\frac{1}{2\pi} \int_0^\infty d\omega \left(\operatorname{Im} \frac{1}{\varepsilon(k, \omega)} \right) + \left(\frac{2\pi n e^2}{k^2} \right) \right]. \quad (9.16)$$

14.10.* Spin and Statistics

Elementary particles with half-integral spins obey Fermi statistics and those with integral spins follow Bose statistics. Electrons and nucleons belong to the former class, and photons and π- and K-mesons to the latter. The statistics of elementary particles can be determined based on the selection rules associated with their reactions, which are described by the relevant Hamiltonians (*14–16*).

Before discussing the general case, let us consider the simplest field due to the Hamiltonian

$$H = \tfrac{1}{2}\{p^2 + q^2\}. \quad (10.1)$$

This is the Hamiltonian for a simple harmonic oscillator. The momentum p and the coordinate q are related to each other through the equation of motion $p = \dot{q}$, where the dot means a time derivative. By applying the transformation

$$\begin{aligned} a &= \frac{1}{\sqrt{2}}(q + ip), \\ a^+ &= \frac{1}{\sqrt{2}}(q - ip), \end{aligned} \quad (10.2)$$

The Hamiltonian is rewritten as follows:

$$\begin{aligned} H &= \tfrac{1}{2}(a^+ a + a a^+) \\ &\equiv \hat{N}, \end{aligned} \quad (10.3)$$

where we have introduced a new operator $\hat{N}$ (for the number of particles).

The operators a and a^+ follow the equation of motion

$$\dot{A} = i[H, A], \quad (10.4)$$

which yields

$$a = [a, \hat{N}], \qquad a^+ = -[a^+, \hat{N}]. \quad (10.5)$$

We find nontrivial eigenvalues:

$$E_n = E_0 + n, \quad (10.6)$$

where E_0 is positive and n is an integer or zero. In the representation that diagonalizes H, only the elements $a_{n, n+1}$ and $a^+_{n-1, n}$ are not zero and

$$a_{n, n+1} = \begin{cases} (2E_0 + n)^{1/2}, & \text{even } n, \\ (n + 1)^{1/2}, & \text{odd } n. \end{cases} \quad (10.7)$$

Because of Eq. (10.7), a and a^+ do not commute. If $E_0 = \frac{1}{2}$, the two cases in Eq. (10.7) coincide, and we obtain commutation relations such as

$$[a, a^+] = 1. \quad (10.8)$$

The Hamiltonian of Eq. (10.3) is symmetric in a and a^+. One might question whether or not this symmetry property has something to do with the above boson commutation relation. This question can be answered by considering an antisymmetric Hamiltonian:

$$\begin{aligned} H &= \tfrac{1}{2}\{a^+ a - a a^+\} \\ &\equiv \hat{N}. \end{aligned} \quad (10.9)$$

Using the same fundamental relation (10.4), for this case we obtain

$$[a, a^+]_+ = 1, \qquad [a, a]_+ = 0, \quad (10.10)$$

that is, anticommutation relations for fermions.

The relations between spin and statistics must in general be discussed relativistically. A detailed account is, of course, beyond the scope of this book, but we can develop treatments similar to the above simple examples (*14–16*).

In relativistic descriptions, the creation and annihilation operators are introduced not only for particles but also for antiparticles. Using a, a^+ for the former and b, b^+ for the latter, we consider a covariant Hamiltonian:

$$H = \sum_{\mathbf{k}} |E_k| (a_k{}^+ a_k - b_k b_k{}^+). \tag{10.11}$$

This Hamiltonian can be decomposed into two parts:

$$H = H_s + H_a, \tag{10.12}$$

where H_s represents the symmetric part,

$$H_s = \tfrac{1}{2} \sum_{\mathbf{k}} |E_k| [(a_k{}^+ a_k + a_k a_k{}^+) - (b_k{}^+ b_k + b_k b_k{}^+)], \tag{10.13}$$

and H_a the antisymmetric part,

$$H_a = \tfrac{1}{2} \sum_{\mathbf{k}} |E_k| [(a_k{}^+ a_k - a_k{}^+ a_k) - (b_k{}^+ b_k - b_k b_k{}^+)]. \tag{10.14}$$

The symmetric Hamiltonian H_s can be shown to correspond to a quantized field of particles with half-integral spins following Bose statistics (or its generalization called para-Bose statistics). However, H_s has energies in the range from $-\infty$ to ∞, and since the energies should be positive, one cannot accept it. On the other hand, if one adopts quantization based on the Fermi (or para-Fermi) statistics, the antisymmetrized Hamiltonian H_a should be used: This Hamiltonian has nonnegative energy eigenvalues and is acceptable. Thus, it is concluded that particles of half-integral spin obey Fermi statistics.

The statistics that enters in the most general discussion of the present subject is the parastatistics introduced in Section 4.9. This is the sense in which we used the term "para" in the above paragraph. If we adopt parastatistics, the number operator can be defined by

$$\hat{N}_k = \tfrac{1}{2}(a_k{}^+ a_k \mp a_k a_k{}^+) \pm \tfrac{1}{2}\hat{p}, \tag{10.15}$$

(upper signs for Fermi, lower sign for Bose case) where $\hat{p}$ takes on eigenvalues 0, 1, ..., p or 0, 1, ..., ∞, the former case corresponding to para-Fermi statistics and the latter to para-Bose statistics.

If the original Hamiltonian is not like Eq. (10.11) but is given by

$$H = \sum_{\mathbf{k}} |E_k| (a_k{}^+ a_k + b_k b_k{}^+), \tag{10.16}$$

we arrive at the reverse situation. One can again introduce symmetrized and antisymmetrized Hamiltonians as in Eqs. (10.13) and (10.14), with the difference that the minus sign is changed into a plus sign for H_s. The Hamiltonian of Eq. (10.16) describes the particles of integral spins. Based on these Hamiltonians and also on the causality principle, it can be concluded that the particles obey Bose statistics.

In Eq. (10.15), we have indicated the possibility of having various eigenvalues for $\hat{p}$. However, all the known elementary particles belong to the cases of $p = 1$ or ∞. It is still a mystery why nature prefers these particular choices, which result in the standard Fermi and Bose statistics. Therefore, efforts have been made to search for new particles that correspond to $p > 1$. There might possibly be some parafermions corresponding to $p = 3$, but there is no conclusive evidence yet.

Problems

1. Show that the operators defined by

$$J_1 = \tfrac{1}{4}i(a^+a^+ - aa), \qquad J_2 = \tfrac{1}{4}(a^+a^+ + aa), \qquad J_3 = \tfrac{1}{4}(a^+a + aa^+)$$

satisfy the following commutation relations:

$$[J_1, J_2] = -iJ_3, \qquad [J_2, J_3] = iJ_1, \qquad [J_3, J_1] = iJ_2.$$

(These relations are for the generators of the three-dimensional Lorentz group.) The operators are assumed to satisfy the usual commutation relations such as $[a, a^+] = 1$.

2. For operators satisfying the anticommutation relations

$$[a, a^+]_+ = 1, \qquad [a, a]_+ = 0$$

define

$$J_1 = \frac{a + a^+}{2}, \qquad J_2 = \frac{a^+ - a}{2i}, \qquad J_3 = \frac{a^+a - aa^+}{2},$$

and find the following relations (for the generators of the three-dimensional orthogonal group):

$$[J_1, J_2] = iJ_3, \qquad [J_2, J_3] = iJ_1, \qquad [J_3, J_1] = iJ_2.$$

3. Show that, to the first-order exchange diagrams, the grand partition function of an electron gas is given by

$$\log \Xi = \frac{2V\beta}{15\pi^2} k_0{}^5 \left(1 + \frac{15}{4\pi a_0 k_0}\right).$$

Here, a_0 is the Bohr radius and k_0 is the Fermi momentum [see Eqs. (5.8) and (5.10)].

4. From the above result, find that the density is given by

$$n = \frac{{k_0}^3}{3\pi^2}\left[1 + \frac{3}{\pi}\left(\frac{4}{9\pi}\right)^{1/3} r_s\right],$$

where $r_s = (9\pi/4)^{1/3}(k_0 a_0)^{-1}$.

5. Using the above grand partition function, evaluate the internal energy and compare the result with Eqs. (5.9).

GENERAL REFERENCES

N. H. March, W. H. Young, and S. Sampanthar, "The Many-Body Problem in Quantum Mechanics." Cambridge Univ. Press, Cambridge, 1967.
C. Kittel, "Quantum Theory of Solids." Wiley, New York, 1963.

REFERENCES

1. N. N. Bogoliubov, *J. Phys. (U.S.S.R.)* **11**, 23 (1947).
2. O. Penrose and L. Onsager, *Phys. Rev.* **104**, 576 (1956).
3. R. P. Feynman and M. Cohen, *Phys. Rev.* **102**, 1189 (1956).
4. T. D. Lee, K. Huang, and C. N. Yang, *Phys. Rev.* **106**, 1135 (1957).
5. A. Isihara and D. D. H. Yee, *Phys. Rev.* A **136**, 618 (1964); C. J. Nisteruk and A. Isihara, *Phys. Rev.* **154**, 150 (1967).
6. D. C. Burnham, F. C. Brown, and R. S. Knox, *Phys. Rev.* **119**, 1560 (1960).
7. H. Fröhlich, H. Pelzer, and S. Zienau, *Phil. Mag.* **41**, 221 (1950).
8. J. Bardeen and J. R. Schrieffer, *in* "Progress in Low Temperature Physics," Vol. 3, p. 170. North-Holland, Amsterdam, 1961.
9. G. Placzek, *Phys. Rev.* **86**, 377 (1952).
10. P. Nozier and D. Pines, *Phys. Rev.* **111**, 442 (1958).
11. U. Fano, *Phys. Rev.* **103**, 1202 (1956).
12. P. Nozier and D. Pines, *Nuovo Cimento* **9**, 470 (1958).
13. H. Ehrenreich and M. Cohen, *Phys. Rev.* **115**, 786 (1959).
14. E. P. Wigner, *Phys. Rev.* **77**, 711 (1950).
15. H. S. Green, *Phys. Rev.* **90**, 270 (1953).
16. Y. Ohnuki and S. Kamefuchi, *Phys. Rev.* **170**, 1279 (1968).

15

Green's Functions

We have already discussed in Chapters 10 and 11 the Green's function solution of the Bloch equation for the density matrix. In this chapter, Green's function methods will be investigated in the language of second quantization. The methods of the preceding chapter are effective when the interactions are weak or when the Hamiltonian can be reduced to a diagonal form. The methods in this chapter are effective in the general case. We shall indicate some basic properties of the Green's functions and see how the poles of one-particle Green's functions are related to quasi particles. We shall find how some of the important results in the previous chapters can be obtained by our new methods. Finally, we conclude this chapter, and the book, with a discussion of the fascinating phenomenon of superconductivity.

15.1. Temperature Green's Function

As discussed in Chapter 10, we can apply quantum field theoretical methods to statistical mechanics by introducing a reciprocal thermal energy variable β in place of time. Thus, we shall talk about creation and annihilation of particles at a certain value of β. However, when β means a variable, rather than a given reciprocal thermal energy $1/kT$, we shall use notations such as β' or β''.

The propagation of a particle from β'' to β' can be expressed by the single-particle *temperature Green's function* defined by (*1–7*)

$$G(\mathbf{k};\beta',\beta'') = -\langle Ta_{\mathbf{k}}(\beta')a_{\mathbf{k}}{}^{+}(\beta'')\rangle. \tag{1.1}$$

Here, $a_{\mathbf{k}}(\beta')$ and $a_{\mathbf{k}}{}^{+}(\beta'')$ are annihilation and creation operators in the "Heisenberg representation,"

$$\begin{aligned} a_{\mathbf{k}}(\beta) &= e^{\beta\mathcal{H}}a_{\mathbf{k}}e^{-\beta\mathcal{H}}, \\ a_{\mathbf{k}}{}^{+}(\beta) &= e^{\beta\mathcal{H}}a_{\mathbf{k}}{}^{+}e^{-\beta\mathcal{H}}, \end{aligned} \tag{1.2}$$

and T is "the Wick time-ordering *operator*." Namely, T orders the permutations of the factors in the bracket in such a way that the β variables increase from right to left. For bosons, T is identical with Dyson's time-ordering operator. In the case of fermions, an odd permutation introduces a minus sign. Thus,

$$G(\mathbf{k};\beta',\beta'') = \begin{cases} -\langle a_{\mathbf{k}}(\beta')a_{\mathbf{k}}{}^{+}(\beta'')\rangle, & \beta' > \beta'' \\ \mp\langle a_{\mathbf{k}}{}^{+}(\beta'')a_{\mathbf{k}}(\beta')\rangle, & \beta' < \beta'' \end{cases} \tag{1.3}$$

(minus for bosons, plus for fermions), where $\langle\cdots\rangle$ is a statistical average in the grand ensemble:

$$\langle\cdots\rangle = \frac{\operatorname{tr}(\cdots e^{-\beta\mathcal{H}})}{\operatorname{tr} e^{-\beta\mathcal{H}}}, \tag{1.4}$$

$$\mathcal{H} = H - \mu N, \qquad e^{-\beta\Omega} = \operatorname{tr} e^{-\beta\mathcal{H}}.$$

The trace means the operation of taking the sum of all diagonal elements of the matrix, where the sum extends over both the number of particles and all possible states with a given number of particles. Thus, Ω is the grand potential and μ is the chemical potential.

The reason why G is called a Green's function may be understood by observing its "equation of motion." We note that

$$\frac{\partial a_{\mathbf{k}}(\beta)}{\partial\beta} = [\mathcal{H}(\beta), a_{\mathbf{k}}(\beta)]. \tag{1.5}$$

Here,

$$\begin{aligned} \mathcal{H}(\beta) &= e^{\beta\mathcal{H}}\mathcal{H}e^{-\beta\mathcal{H}} = \mathcal{H} \\ &= \sum_{\mathbf{k}}(\varepsilon_k{}^{0} - \mu)a_{\mathbf{k}}{}^{+}(\beta)a_{\mathbf{k}}(\beta) + \frac{1}{2V}\sum(\mathbf{kq}|u|\mathbf{k'q'})a_{\mathbf{k}}{}^{+}(\beta) \\ &\quad \times a_{\mathbf{q}}{}^{+}(\beta)a_{\mathbf{q}'}(\beta)a_{\mathbf{k}'}(\beta). \end{aligned} \tag{1.6}$$

Thus,

$$\frac{\partial a_{\mathbf{k}}(\beta)}{\partial\beta} = -(\varepsilon_{\mathbf{k}}{}^{0} - \mu)a_{\mathbf{k}}(\beta) - \frac{1}{V}\sum(\mathbf{kq}|u|\mathbf{k'q'})a_{\mathbf{q}}{}^{+}(\beta)a_{\mathbf{q}'}(\beta)a_{\mathbf{k}'}(\beta). \tag{1.7}$$

We note that $G(\mathbf{k};\beta',\beta'')$ is discontinuous at $\beta' = \beta''$ because

$$\begin{aligned} G(\mathbf{k};\beta'+0,\beta') - G(\mathbf{k};\beta'-0,\beta') &= -\langle a_{\mathbf{k}}(\beta')a_{\mathbf{k}}{}^{+}(\beta') \mp a_{\mathbf{k}}{}^{+}(\beta')a_{\mathbf{k}}(\beta')\rangle \\ &= -1. \end{aligned} \tag{1.8}$$

Keeping this discontinuity in mind, we obtain

$$\begin{aligned} &\frac{\partial G(\mathbf{k};\beta',\beta'')}{\partial\beta'} + (\varepsilon_k{}^{0} - \mu)G(\mathbf{k};\beta',\beta'') \\ &\quad = -\delta(\beta'-\beta'') + \frac{1}{V}\sum_{\mathbf{q},\mathbf{k}',\mathbf{q}'}(\mathbf{kq}|u|\mathbf{k'q'})\langle Ta_{\mathbf{q}}{}^{+}(\beta')a_{\mathbf{q}'}(\beta')a_{\mathbf{k}'}(\beta')a_{\mathbf{k}}{}^{+}(\beta'')\rangle. \end{aligned} \tag{1.9}$$

The last term can be expressed by the two-particle Green's function [see Eq. (8.7)]. For free particles, Eq. (1.9) reduces to

$$\frac{\partial G_0(\mathbf{k};\beta',\beta'')}{\partial\beta'} + (\varepsilon_k{}^{0} - \mu)G_0(\mathbf{k};\beta',\beta'') = -\delta(\beta'-\beta''), \tag{1.10}$$

which is simply the equation for the Green's function G_0 associated with the Bloch-type differential equation.

The one-particle temperature Green's function can be defined in the $\mathbf{r}$–β space as follows:

$$G(\mathbf{r}\beta';\mathbf{r}'\beta'') = -\langle T(\Psi(\mathbf{r}\beta')\Psi^{+}(\mathbf{r}'\beta'')\rangle, \tag{1.11}$$

where $\Psi(\mathbf{r},\beta)$ and $\Psi^{+}(\mathbf{r},\beta)$ are annihilation and creation operators at $\mathbf{r}$ in the Heisenberg representation:

$$\begin{aligned} \Psi(\mathbf{r}\beta') &= (\exp\beta'\mathcal{H})\psi(\mathbf{r})(\exp-\beta'\mathcal{H}), \\ \Psi^{+}(\mathbf{r}'\beta'') &= (\exp\beta''\mathcal{H})\psi(\mathbf{r}')(\exp-\beta''\mathcal{H}). \end{aligned}$$

15.2. Properties of the Green's Function

Here, we study some basic properties of the one-body temperature Green's function. For this purpose, we use the representation in which $\mathcal{H}$ is diagonal:

$$\mathcal{H}|l\rangle = (E_l - N_l\mu)|l\rangle. \tag{2.1}$$

(*i*) First, we show that

$$G(\mathbf{k};\beta',\beta'') = G(\mathbf{k};\beta'-\beta''). \tag{2.2}$$

In other words, G is a function of the difference $\beta'-\beta''$.

This property can be proved by just writing down the Heisenberg operators in the definition, Eq. (1.1), of the Green's function and then by making use of the cyclic invariance of the trace. In the representation corresponding to Eq. (2.1), the proof for fermions is as follows. (For bosons, similar steps can be taken.) Consider the case $\beta' > \beta''$. Equation (1.1) gives

$$\begin{aligned} G(\mathbf{k}; \beta', \beta'') &= -\mathrm{tr}\{[\exp \beta(\Omega - \mathscr{H})](\exp \beta'\mathscr{H})a_{\mathbf{k}} \\ &\quad \times (\exp -\beta'\mathscr{H})(\exp \beta''\mathscr{H})a_{\mathbf{k}}{}^{+} (\exp -\beta''\mathscr{H})\} \\ &= -\sum_{l,m} z^{N_l}[\exp \beta(\Omega - E_l)][\exp -(\beta' - \beta'')(E_m - E_l - \mu)] \\ &\quad \times \langle l|a_{\mathbf{k}}|m\rangle\langle m|a_{\mathbf{k}}{}^{+}|l\rangle, \qquad z = e^{\beta\mu}, \end{aligned} \tag{2.3}$$

where the chemical potential μ appears because $a_{\mathbf{k}}{}^{+}$ brings one particle with momentum $\mathbf{k}$ from the state $|l\rangle$ to the state $|m\rangle$. Equation (2.3) proves Property (*i*). For $\beta' < \beta''$, Eq. (2.3) is replaced by

$$\begin{aligned} G(\mathbf{k}; \beta', \beta'') &= \mathrm{tr}\{[\exp \beta(\Omega - \mathscr{H})](\exp \beta''\mathscr{H})a_{\mathbf{k}}{}^{+} \\ &\quad \times (\exp -\beta''\mathscr{H})(\exp \beta'\mathscr{H})a_{\mathbf{k}}(\exp -\beta'\mathscr{H})\} \\ &= \sum_{l,m} z^{N_l}[\exp \beta(\Omega - E_l)][\exp -(\beta'' - \beta')(E_m - E_l + \mu)] \\ &\quad \times \langle l|a_{\mathbf{k}}{}^{+}|m\rangle\langle m|a_{\mathbf{k}}|l\rangle. \end{aligned} \tag{2.4}$$

From Eq. (2.2), we shall use the notation $G(\mathbf{k}, \alpha)$ for the Green's function.

(*ii*) The second property is

$$G(\mathbf{k}, \beta + \alpha) = \mp G(\mathbf{k}, \alpha) \tag{2.5}$$

(the minus sign is for fermions, the plus sign for bosons), where

$$\alpha = \beta' - \beta''. \tag{2.6}$$

The proof for $\beta > \alpha > 0$ may be made by using Eq. (2.3) and the definition of the Green's function. If $-\beta < \alpha < 0$, then $0 < \alpha + \beta < \beta$. Therefore, we use Eq. (2.4) and then Eq. (2.3) to find the property.

(*iii*) The third property follows from Eq. (2.5):

$$G(\mathbf{k}, 2\beta + \alpha) = G(\mathbf{k}, \alpha). \tag{2.7}$$

Based on this property, one can introduce a Fourier expansion

$$G(\mathbf{k}, \alpha) = \frac{1}{\beta}\sum_{n} \tilde{G}\left(\mathbf{k}, \frac{\pi i n}{\beta}\right)e^{-\pi i n\alpha/\beta}. \tag{2.8}$$

The Fourier transform is given by

$$\tilde{G}\left(\mathbf{k}, \frac{2\pi i n}{2\beta}\right) = \frac{1}{2}\left[\int_{-\beta}^{0} + \int_{0}^{\beta}\right] G(k, \alpha)e^{2\pi i n\alpha/2\beta}\, d\alpha, \tag{2.9}$$

$$= \frac{1}{2}(1 \mp e^{i\pi n})\int_{0}^{\beta} G(k, \alpha)e^{2\pi i n\alpha/2\beta}\, d\alpha, \tag{2.10}$$

from Eq. (2.5). Thus, $\tilde{G}$ vanishes for even n in the case of fermions and for odd n in the case of bosons. Accordingly, we define

$$\omega_n \begin{cases} = \dfrac{(2n+1)\pi}{\beta} & \text{(fermions)}, \\ = \dfrac{2n\pi}{\beta} & \text{(bosons)}, \end{cases} \qquad (n = 0, \pm 1, \pm 2, \ldots). \tag{2.11}$$

In terms of ω_n, the Fourier-transformed Green's function is given by

$$\tilde{G}(\mathbf{k}, i\omega_n) = \int_{0}^{\beta} G(\mathbf{k}, \alpha)e^{i\omega_n\alpha}\, d\alpha. \tag{2.12}$$

In the next section, we shall study the statistical averages of products of operators and obtain an explicit expression for the free-particle Green's function.

15.3. Contraction

We shall find shortly that the grand partition function and the Green's functions of interacting systems can be obtained by taking averages in terms of $\rho_0 = e^{-\beta\mathscr{H}_0}$ instead of $\rho = e^{-\beta\mathscr{H}}$. Here, $\mathscr{H}_0 = H_0 - \mu N$, H_0 is the unperturbed Hamiltonian, N is the operator that produces the number of particles, and μ is the chemical potential. We shall assume that ρ_0 is diagonal. Let us study the method of evaluating averages of products of creation and annihilation operators in the grand ensemble with ρ_0 as the weight factor.

In general, the averages of these products are zero except for the cases when the number of annihilation operators is equal to that of the creation operators belonging to the same state.

The statistical average of a pair of operators is called a *contraction*. The contractions $\langle aa\rangle_0$ and $\langle a^{+}a^{+}\rangle_0$ are zero, where $\langle\cdots\rangle_0$ is the average in terms of $e^{-\beta\mathscr{H}_0}$. The only nonvanishing contractions are

$$\langle a_{\mathbf{k}}{}^{+}a_{\mathbf{k}}\rangle_0 = \langle N_{\mathbf{k}}\rangle_0 = \frac{1}{z^{-1}(\exp + \beta k^2) - \theta}, \tag{3.1}$$

$$\langle a_{\mathbf{k}}\, a_{\mathbf{k}}{}^{+}\rangle_0 = \langle 1 + \theta N_k\rangle_0 = \frac{1}{1 - \theta z(\exp - \beta k^2)}, \tag{3.2}$$

where $\theta = -1$ for fermions and $\theta = +1$ for bosons. Note that these relations are due to the general commutation relation

$$A_{\mathbf{k}} A_{\mathbf{s}} - \theta A_{\mathbf{s}} A_{\mathbf{k}} = \delta(\mathbf{k}, \mathbf{s}),$$

$$\delta(\mathbf{k}, \mathbf{s}) = \begin{cases} \delta_{\mathbf{k},\mathbf{s}} & \text{(fermions)} \\ \pm\delta_{\mathbf{k},\mathbf{s}} & \text{(bosons, plus for } A_{\mathbf{k}} = a_{\mathbf{k}}\text{, minus for } A_{\mathbf{k}} = a_{\mathbf{k}}^{+}\text{).} \end{cases} \tag{3.3}$$

Indeed,

$$\langle A_{\mathbf{k}} A_{\mathbf{s}} \rangle_0 = \delta(\mathbf{k}, \mathbf{s}) + \theta \langle A_{\mathbf{s}} A_{\mathbf{k}} \rangle_0 , \tag{3.4a}$$

$$= \delta(\mathbf{k}, \mathbf{s}) + \theta \operatorname{tr}(A_{\mathbf{s}} A_{\mathbf{k}} \exp - \beta \mathscr{H}_0), \tag{3.4b}$$

$$= \delta(\mathbf{k}, \mathbf{s}) + \theta \operatorname{tr}[A_{\mathbf{s}}(\exp -\beta \mathscr{H}_0) A_{\mathbf{k}}](z \exp -\beta\varepsilon_k^{0})^{\pm 1} \tag{3.4c}$$

$$= \delta(\mathbf{k}, \mathbf{s}) + \theta \langle A_{\mathbf{k}} A_{\mathbf{s}} \rangle_0 (z \exp - \beta\varepsilon_k^{0})^{\pm 1}. \tag{3.4d}$$

Thus,

$$\langle A_{\mathbf{k}} A_{\mathbf{s}} \rangle_0 = \frac{\delta(\mathbf{k}, \mathbf{s})}{[1 - \theta(z \exp -\beta\varepsilon_k^{0})^{\pm 1}]}, \tag{3.5}$$

where the plus and minus signs correspond respectively to $A_{\mathbf{k}} = a_{\mathbf{k}}$ and $A_{\mathbf{k}} = a_{\mathbf{k}}^{+}$.

The last equality (3.4d) is based on the invariance of the trace under a cyclic permutation, and the third one (3.4c) is proved as follows. Define

$$a_{\mathbf{k}}(\beta) = e^{\beta \mathscr{H}_0} a_{\mathbf{k}} e^{-\beta \mathscr{H}_0}. \tag{3.6}$$

$$\mathscr{H}_0 = \sum_{\mathbf{k}} (\varepsilon_k^{0} - \mu) a_{\mathbf{k}}^{+} a_{\mathbf{k}}. \tag{3.7}$$

Then,

$$\frac{\partial a_{\mathbf{k}}(\beta)}{\partial \beta} = e^{\beta \mathscr{H}_0}(\mathscr{H}_0 a_{\mathbf{k}} - a_{\mathbf{k}} \mathscr{H}_0) e^{-\beta \mathscr{H}_0}, \tag{3.8}$$

and we find

$$\frac{\partial a_{\mathbf{k}}(\beta)}{\partial \beta} = -(\varepsilon_k^{0} - \mu) a_{\mathbf{k}}(\beta), \tag{3.9}$$

and, upon integration,

$$\begin{aligned} a_{\mathbf{k}}(\beta) &= a_{\mathbf{k}}(z \exp -\beta\varepsilon_k^{0}), \\ a_{\mathbf{k}}(\exp -\beta \mathscr{H}_0) &= (\exp -\beta \mathscr{H}_0) a_{\mathbf{k}}(z \exp -\beta\varepsilon_k^{0}). \end{aligned} \tag{3.10}$$

Similarly, we obtain

$$\begin{aligned} a_{\mathbf{k}}^{+}(\beta) &= a_{\mathbf{k}}^{+}(z \exp -\beta\varepsilon_k^{0})^{-1}, \\ a_{\mathbf{k}}^{+}(\exp -\beta \mathscr{H}_0) &= (\exp -\beta \mathscr{H}_0) a_{\mathbf{k}}^{+}(z \exp -\beta\varepsilon_k^{0})^{-1}. \end{aligned} \tag{3.11}$$

In the evaluation of the contraction $\langle A_{\mathbf{k}} A_{\mathbf{s}} \rangle_0$ in Eqs. (3.4), it is important to observe the following rules (for fermions):

(1) The operator $A_{\mathbf{k}}$ is transferred to the extreme right in the operator product.
(2) A permutation of a pair of operators introduces a minus sign.
(3) $A_{\mathbf{k}}$ is brought back to the left in the operator product.

These rules can be followed if there are many operators. For instance, a product of four-fermion operators A_i can be expanded as follows:

$$\begin{aligned} [1 + (z \exp -\beta\varepsilon_k^{0})^{\pm 1}] \langle A_1 A_2 A_3 A_4 \rangle_0 = [&\delta(1, 2) \langle A_3 A_4 \rangle_0 - \delta(1, 3) \langle A_2 A_4 \rangle_0 \\ &+ \delta(1, 4) \langle A_2 A_3 \rangle_0]. \end{aligned} \tag{3.12}$$

By applying Eq. (3.5), this can be rewritten

$$\begin{aligned} \langle A_1 A_2 A_3 A_4 \rangle_0 = &\langle A_1 A_2 \rangle_0 \langle A_3 A_4 \rangle_0 \\ &- \langle A_1 A_3 \rangle_0 \langle A_2 A_4 \rangle_0 + \langle A_1 A_4 \rangle_0 \langle A_2 A_3 \rangle_0 . \end{aligned} \tag{3.13}$$

Thus, repeating the same process, we arrive at a formula (for fermions)

$$\langle A_1 A_2 \cdots A_{2n} \rangle_0 = \sum_P (-)^P \langle A_{i_1} A_{i_2} \rangle_0 \langle A_{i_3} A_{i_4} \rangle_0 \cdots \langle A_{i_{2n-1}} A_{i_{2n}} \rangle_0 . \tag{3.14}$$

Here, a pair of operators in each contraction on the right-hand side is ordered such that $i_{n-1} < i_n$ and the entire product is arranged such that $i_1 < i_3 < \cdots < i_{2n-1}$. If the sequence $i_1, i_2, \ldots, i_{2n}$ is obtained from the original sequence $1, 2, \ldots, 2n$ by an even permutation, P is even. If the sequence is given by odd permutations, P is odd. In other words, the terms on the right-hand side carry a plus or a minus sign, depending upon even or odd permutations.

Equation (3.14) represents an important reduction formula for a contraction of a product of operators. The rules involved in this formula are:

(1) Make all possible pairings from the original product and sum over all such pairing sets.
(2) Attach a sign to each set of pairings according to whether it is obtained by an even permutation or an odd permutation from the original product.
(3) In the case of bosons, there will be no minus sign for permutations.

Let us now evaluate the free-particle temperature Green's function. For fermions, the first of Eqs. (1.3) gives

$$G_0(\mathbf{k}, \alpha) = -\langle a_{\mathbf{k}} a_{\mathbf{k}}^{+} \rangle_0 [\exp -\alpha(k^2 - \mu)], \tag{3.15}$$

where $\alpha = (\beta' - \beta'') > 0$. Taking the Fourier transform defined by Eq. (2.12), we find

$$\begin{aligned}\tilde{G}_0(\mathbf{k}, i\omega_n) &= -\int_0^\beta \langle a_\mathbf{k} a_\mathbf{k}^+ \rangle_0 \exp[-\alpha(k^2 - \mu) + i\omega_n \alpha]\, d\alpha \\ &= \langle a_\mathbf{k} a_\mathbf{k}^+ \rangle_0 \frac{[\exp -\beta(k^2 - \mu)] + 1}{i\omega_n - k^2 + \mu} \\ &= \frac{1}{i\omega_n - k^2 + \mu},\end{aligned} \tag{3.16}$$

where we have used Eqs. (2.11) and (3.2). The Fourier inverse of Eq. (3.16) is

$$G_0(\mathbf{k}, \beta' - \beta'') = \frac{1}{\beta} \sum_n \frac{\exp[-i\omega_n(\beta' - \beta'')]}{i\omega_n - k^2 + \mu}. \tag{3.17}$$

We obtain similar expressions for bosons. For bosons, ω_n can be zero [see Eq. (2.11)]. The corresponding Green's function $\tilde{G}_0$ is

$$\tilde{G}_0(\mathbf{k}, 0) = (\mu - k^2)^{-1}. \tag{3.18}$$

The right-hand side is infinite for $\mathbf{k} = 0$ and $\mu = 0$. The Bose condensation corresponds to such a divergence, which takes place at a finite temperature. At this "critical" temperature, the number of particles in the state $\mathbf{k} = 0$ starts increasing to a large macroscopic value.

15.4. Perturbation Calculation of the Grand Partition Function

Let us consider evaluating the grand partition function defined by

$$\begin{aligned}\Xi &= \operatorname{tr} e^{-\beta(H - \mu N)} \\ &= e^{-\beta\Omega}\end{aligned} \tag{4.1}$$

for an H of the following structure:

$$H = H_0 + H_1, \tag{4.2}$$

where

$$\begin{aligned}H_0 &= \sum \varepsilon_k^{\,0} a_\mathbf{k}^+ a_\mathbf{k}, \\ H_1 &= \frac{1}{2V} \sum_{\mathbf{q}\mathbf{k}\mathbf{k}'} u(q) a_{\mathbf{k}+\mathbf{q}}^+ a_{\mathbf{k}'+\mathbf{q}}^+ a_{\mathbf{k}'} a_\mathbf{k}.\end{aligned} \tag{4.3}$$

We define

$$\begin{aligned}\mathscr{H}_0 &= H_0 - \mu N \\ &= \sum_\mathbf{k} (\varepsilon_k^{\,0} - \mu) a_\mathbf{k}^+ a_\mathbf{k}.\end{aligned} \tag{4.4}$$

$$\mathscr{H} = H - \mu N. \tag{4.5}$$

Then,

$$\Xi = \operatorname{tr}[e^{-\beta\mathscr{H}_0} R(\beta)] = e^{-\beta\Omega_0} \langle R(\beta) \rangle_0, \tag{4.6}$$

$$\frac{\Xi}{\Xi_0} = \operatorname{tr} \rho_0 R = \langle R \rangle_0, \tag{4.7}$$

where Ω_0 is the grand potential corresponding to $\mathscr{H}_0$ and

$$R(\beta) = e^{\beta\mathscr{H}_0} e^{-\beta\mathscr{H}}, \tag{4.8}$$

$$\rho_0 = e^{-\beta\mathscr{H}_0},$$

$$\langle \cdots \rangle_0 = \frac{\operatorname{tr}(\cdots e^{-\beta\mathscr{H}_0})}{\operatorname{tr} e^{-\beta\mathscr{H}_0}}. \tag{4.9}$$

By definition,

$$\frac{\partial R(\beta)}{\partial \beta} = -e^{\beta\mathscr{H}_0}(\mathscr{H} - \mathscr{H}_0) e^{-\beta\mathscr{H}}.$$

Thus,

$$\frac{\partial R(\beta)}{\partial \beta} = -H_1(\beta) R(\beta), \tag{4.10}$$

where

$$H_1(\beta) = e^{\beta\mathscr{H}_0} H_1 e^{-\beta\mathscr{H}_0}. \tag{4.11}$$

Equation (4.10) is Bloch's equation in the interaction representation. With $R(0) = 1$, it is equivalent to the integral equation

$$R(\beta) = 1 - \int_0^\beta H_1(\beta) R(\beta)\, d\beta. \tag{4.12}$$

The iterative solution of this equation is

$$R(\beta) = 1 + \sum_{n=1}^{\infty} (-)^n \int_0^\beta d\beta_1 \int_0^{\beta_1} d\beta_2 \cdots \int_0^{\beta_{n-1}} d\beta_n\, H_1(\beta_1) \cdots H_1(\beta_n). \tag{4.13}$$

Combining this result with Eq. (4.7), we obtain

$$\frac{\Xi}{\Xi_0} = \sum_{n=1}^{\infty} (-)^n \int_0^\beta \int_0^{\beta_1} \cdots \int_0^{\beta_{n-1}} d\beta_1\, d\beta_2 \cdots d\beta_n \langle H_1(\beta_1) H_1(\beta_2) \cdots H_1(\beta_n) \rangle_0 \tag{4.14}$$

$$= 1 + \sum_{n=1}^{\infty} \frac{(-)^n}{n!} \int_0^\beta \cdots \int_0^\beta d\beta_1 \cdots d\beta_n \langle P[H_1(\beta_1) H_1(\beta_2) \cdots H_1(\beta_n)] \rangle_0. \tag{4.15}$$

where P is Dyson's ordering parameter, which requires that the β's be arranged such that they decrease from left to right. Note that, if H_1, or, more generally, a product of operators, includes only even numbers of creation and annihilation operators, then

$$\begin{aligned} T[H_1(\beta_1)H_1(\beta_2)] &= P[H_1(\beta_1)H_1(\beta_2)], \\ T[H_1(\beta_1)H_1(\beta_2)] &= T[H_1(\beta_2)H_1(\beta_1)]. \end{aligned} \tag{4.16}$$

Using Eq. (4.3) on the right-hand side of Eq. (4.15), we find

$$\frac{\Xi}{\Xi_0} = 1 + \sum_{n=1}^{\infty} \frac{(-)^n}{n!(2V)^n} \sum_{\substack{\mathbf{pq} \\ \mathbf{p'q'}}} \prod_{i=1}^{n} (\mathbf{p}_i\mathbf{q}_i|u|\mathbf{p}_i'\mathbf{q}_i')\,\delta(\mathbf{p}_i + \mathbf{q}_i - \mathbf{p}_i' - \mathbf{q}_i') \times \int_0^\beta \cdots \int_0^\beta d\beta_1 \cdots d\beta_n \langle P[a_{\mathbf{p}_i}^+(\beta_1)a_{\mathbf{q}_i}^+(\beta_1)a_{\mathbf{q}_i'}(\beta_1)a_{\mathbf{p}_i'}(\beta_1)\cdots]\rangle_0 . \tag{4.17}$$

One can apply the rules of evaluating contractions to the products on the right-hand sides of operators. It is important to note that the products can break into several closed cycles of operators. For instance, in the first-order term,

$$-\frac{1}{2V}\sum (\mathbf{pq}|u|\mathbf{p'q'}) \int_0^\beta \langle a_{\mathbf{p}}^+(\beta_1)a_{\mathbf{q}}^+(\beta_1)a_{\mathbf{q}'}(\beta_1)a_{\mathbf{p}'}(\beta_1)\rangle_0\, d\beta_1, \tag{4.18}$$

the integrand is split into two parts:

$$\begin{aligned} &\langle a_{\mathbf{p}}^+(\beta_1)a_{\mathbf{p}'}(\beta_1)\rangle_0\langle a_{\mathbf{q}}^+(\beta_1)a_{\mathbf{q}'}(\beta_1)\rangle_0 \\ &\quad + \theta\langle a_{\mathbf{p}}^+(\beta_1)a_{\mathbf{q}'}(\beta_1)\rangle_0\langle a_{\mathbf{q}}^+(\beta_1)a_{\mathbf{p}'}(\beta_1)\rangle_0 . \end{aligned} \tag{4.19}$$

The first term contributes only when $\mathbf{p} = \mathbf{p}'$ and $\mathbf{q} = \mathbf{q}'$, and the second term when $\mathbf{p} = \mathbf{q}'$ and $\mathbf{q} = \mathbf{p}'$. Diagrammatically, these two terms may be represented, respectively, by two circles and by one closed loop of two vertices.

The term *vertex* means a point in β-space at which a pair of creation and annihilation operators belonging to the same state appears. In the above example, both terms have two vertices. The first is composed of two separate graphs which are connected by an interaction line. The second yields in Eq. (4.18) a closed-loop configuration with one "internal" interaction line. The end points of these interactions are vertices.

Higher-order terms can be represented similarly by diagrams. Thus, according to our considerations in Chapter 10, or by applying the cumulant theorem, we conclude that $\log(\Xi/\Xi_0)$ is represented by a sum of connected graphs.

Moreover, we can use the free-particle Green's function of Eq. (3.17),

$$\begin{aligned} G_0(\mathbf{k}, \beta' - \beta'') &= -\langle T a_{\mathbf{k}}(\beta')a_{\mathbf{k}}^+(\beta'')\rangle_0 \\ &= \frac{1}{\beta}\sum_n \frac{\exp[-i\omega_n(\beta' - \beta'')]}{i\omega_n - \varepsilon_k^0 + \mu}, \end{aligned} \tag{4.20}$$

in the products of contractions. Furthermore, the sign of a connected graph is determined by $(-)^{n+c}$, where n represents the order (number of interaction lines) and c is the number of component graphs in the diagram. Thus, we obtain for fermions

$$\log\frac{\Xi}{\Xi_0} = \sum_{n=1}^{\infty} \frac{(-)^n}{n!(2V)^n} \sum_c (-)^c \prod_{i=1}^{n} (\mathbf{p}_i\mathbf{q}_i|u|\mathbf{p}_i'\mathbf{q}_i') \times \int_0^\beta \prod G_0(\mathbf{k}; \beta_{i_1}, \beta_{i_2})\, d\beta_1 \cdots d\beta_n , \tag{4.21}$$

where we have expressed the right-hand side in ascending powers in the interaction potential. The second sum is over all connected graphs of the order n, over which the first sum is being performed, and the product of the free-particle Green's functions represents all the free-particle propagations in the connected graphs.

The evaluation of Eq. (4.21) is greatly simplified because of the form of Eq. (4.20). That is, one can perform the integrations over the β_i to arrive at energy conservations of the form

$$\omega_n + \omega_m = \omega_r + \omega_s . \tag{4.22}$$

According to Eq. (2.11), this equation reduces to

$$n + m = r + s. \tag{4.23}$$

Further considerations require specific diagrams and detailed analyses. Since we have already discussed an alternative method in Chapter 10, we shall not repeat the detailed analyses. However, we remark that the grand partition function in the ring diagram approximation is given by

$$\log\frac{\Xi}{\Xi_0} = \frac{1}{2\beta}\sum_{\mathbf{q},l} \{u(q)\lambda_l(q) - \log[1 + u(q)\lambda_l(q)]\}, \tag{4.24}$$

where

$$\lambda_l(q) = -\frac{2}{\beta V}\sum_{\mathbf{k},m} [(i\omega_m - \varepsilon_k^0 + \mu)(i\omega_{m+l} - \varepsilon_{\mathbf{k}+\mathbf{q}}^0 + \mu)]^{-1}, \tag{4.25}$$

in which a weight factor 2 has been introduced to take the spins into account.

The summation in Eq. (4.24) can be carried out for $\beta \to \infty$ by replacing the summation by integration over a variable $m/\beta = y$. The result is determined by the limiting variable $l/\beta = x$

$$\lambda(q, x) = \frac{2}{V}\sum_{\mathbf{k}} \frac{[f^{(0)}(\mathbf{k} + \mathbf{q}) - f^{(0)}(k)](\varepsilon_k^0 - \varepsilon_{\mathbf{k}+\mathbf{q}}^0)}{(\varepsilon_{\mathbf{k}+\mathbf{q}}^0 - \varepsilon_k^0)^2 + (2\pi x)^2} \tag{4.26}$$

to be

$$\log\frac{\Xi}{\Xi_0} = \frac{1}{2}\sum_{\mathbf{q}} \int_{-\infty}^{\infty} \{u(q)\lambda(q, x) - \log[1 + u(q)\lambda(q, x)]\}\, dx, \tag{4.27}$$

where $f^{(0)}$ is the Fermi distribution.

Equation (4.27) agrees with Eq. (7.20) of Chapter 10. Also, dropping the factor 2 due to spins and making use of

$$\frac{1}{V}\sum_{\mathbf{k}} \to \frac{1}{(2\pi)^3}\int d\mathbf{k},$$

one can show easily that Eq. (4.26) reduces to Eq. (8.2) of Chapter 10.

15.5. Interaction Representation and Lehmann Representation

The averaging in terms of $e^{-\beta\mathscr{H}}$ in Eq. (1.11) can be carried out in terms of $e^{-\beta\mathscr{H}_0}$ with $R(\beta)$ as the conversion factor. To see this, let us make use of Eq. (4.8) and its inverse:

$$\begin{aligned} \exp(-\beta'\mathscr{H}) &= (\exp -\beta'\mathscr{H}_0)R(\beta'), \\ \exp \beta'\mathscr{H} &= [R(\beta')]^{-1}\exp \beta'\mathscr{H}_0. \end{aligned} \tag{5.1}$$

The single-particle Green's function becomes

$$\begin{aligned} &G(\mathbf{r}\beta';\mathbf{r}'\beta'') \\ &\quad = -\langle T\Psi(\mathbf{r},\beta')\Psi^+(\mathbf{r}',\beta'')\rangle \\ &\quad = -\frac{\mathrm{tr}\{e^{-\beta\mathscr{H}_0}R(\beta)[R(\beta')]^{-1}\Psi(\mathbf{r},\beta')R(\beta')[R(\beta'')]^{-1}\Psi^+(\mathbf{r}',\beta'')R(\beta'')\}}{\mathrm{tr}[e^{-\beta\mathscr{H}_0}R(\beta)]}, \end{aligned} \tag{5.2}$$

where

$$\Psi(\mathbf{r},\beta') = (\exp \beta'\mathscr{H}_0)\psi(\mathbf{r})(\exp -\beta'\mathscr{H}_0), \quad \text{etc.}$$

It is convenient to express $R(\beta)$ of Eq. (4.13) symbolically by

$$R(\beta) = P\exp\left[-\int_0^\beta H_1(\beta_1)\,d\beta_1\right], \tag{5.3}$$

where the ordered exponential function represents the series of Eq. (4.15). We note that

$$\begin{aligned} R(\beta')[R(\beta'')]^{-1} &= P\exp\left[-\int_{\beta''}^{\beta'} H_1(\beta_1)\,d\beta_1\right] \\ &= R(\beta',\beta''). \end{aligned} \tag{5.4}$$

The new function $R(\beta',\beta'')$ defined by the last equality satisfies the following:

$$\begin{aligned} R(\beta',\beta''') &= R(\beta',\beta'')R(\beta'',\beta'''), \\ R(\beta') &= R(\beta',0). \end{aligned} \tag{5.5}$$

Using these properties and those of the ordering parameter P (or T), we find that the product of the R's in Eq. (5.2) can be arranged such that

$$R(\beta,\beta')R(\beta',\beta'')R(\beta'',0) = R(\beta,0). \tag{5.6}$$

Therefore, we arrive at the Green's function in the interaction representation:

$$\begin{aligned} G(\mathbf{r}\beta';\mathbf{r}'\beta'') &= -\frac{\mathrm{tr}\,T[e^{-\beta\mathscr{H}_0}R(\beta)\Psi(\mathbf{r},\beta')\Psi^+(\mathbf{r}',\beta'')]}{\mathrm{tr}[e^{-\beta\mathscr{H}_0}R(\beta)]} \\ &= -\frac{T\langle R(\beta)\Psi(\mathbf{r},\beta')\Psi^+(\mathbf{r}',\beta'')\rangle_0}{\langle R(\beta)\rangle_0}. \end{aligned} \tag{5.7}$$

Similar expressions hold for higher-order Green's functions, which have more factors of annihilation and creation operators.

The discussion which followed Eq. (4.19) is applicable to the diagram representation of $G(\mathbf{r}\beta',\mathbf{r}'\beta'')$. As in the case of the p.d.f. discussed in Chapter 11, the diagrams for G can be divided into two groups: connected graphs and disconnected graphs. The contribution from the disconnected graphs cancels the denominator of Eq. (5.7). Therefore, we obtain

$$G(\mathbf{r}\beta',\mathbf{r}'\beta'') = -T\langle R(\beta)\Psi(\mathbf{r},\beta')\Psi^+(\mathbf{r}',\beta'')\rangle_{0c}, \tag{5.8}$$

where the subscript "c" indicates "connected."

Let us now discuss what is known as the *Lehmann representation*. By definition and from the properties of the Green's function, $\tilde{G}(\mathbf{k}, i\omega)$ can be represented as follows:

$$\tilde{G}(\mathbf{k}, i\omega) = \int_{-\infty}^{\infty} \frac{\rho(\mathbf{k},x)}{i\omega - x}\,dx, \tag{5.9}$$

where $\rho(\mathbf{k},x)$ satisfies

$$\int \rho(\mathbf{k},x)\,dx = 1. \tag{5.10}$$

The relation (5.9) can be considered to define $\tilde{G}(\mathbf{k}, i\omega)$ for all complex values.

The proof of Eq. (5.9) can be made for fermions by introducing Eq. (2.3) in Eq. (2.12). For $\alpha > 0$,

$$\begin{aligned} \tilde{G}(\mathbf{k}, i\omega_n) &= -\sum_{l,m} z^{N_l}e^{\beta(\Omega-E_l)}\langle l|a_{\mathbf{k}}|m\rangle\langle m|a_{\mathbf{k}}{}^+|l\rangle\int_0^\beta e^{(E_l-E_m+\mu+i\omega_n)\alpha}\,d\alpha \\ &= \sum_{l,m} z^{N_l}e^{\beta(\Omega-E_l)}\frac{e^{\beta(E_l-E_m+\mu)}+1}{i\omega_n-(E_m-E_l-\mu)}\langle l|a_{\mathbf{k}}|m\rangle\langle m|a_{\mathbf{k}}{}^+|l\rangle. \end{aligned} \tag{5.11}$$

Thus, defining

$$\rho(\mathbf{k}, x) = \sum_{l,m} z^{N_l} e^{\beta(\Omega - E_l)} \langle l | a_{\mathbf{k}} | m \rangle \langle m | a_{\mathbf{k}}^{+} | l \rangle (e^{-\beta x} + 1)\, \delta(x - (E_m - E_l - \mu)) \tag{5.12}$$

and using the analytic continuation $\omega_n \to \omega$, we obtain the foregoing expansion. The normalization property of $\rho(\mathbf{k}, x)$ can easily be proved from Eq. (5.12).

For free particles, $E = \sum N_s \varepsilon_s^{0}$. Since the state $|m\rangle$ has one more particle with momentum $\mathbf{k}$ than the state $|l\rangle$ has and because of Eq. (3.2), we obtain

$$\rho_0(\mathbf{k}, x) = \delta(x - \varepsilon_k^{0} + \mu), \tag{5.13}$$

and

$$\tilde{G}_0(\mathbf{k}, i\omega) = (i\omega - \varepsilon_k^{0} + \mu)^{-1}. \tag{5.14}$$

The latter expression agrees with our previous result (3.16).

We remark that the pole of $\tilde{G}_0(\mathbf{k}, i\omega)$ is simply at the energy of the free particle measured from μ. The matrix element $\langle m | a_{\mathbf{k}}^{+} | l \rangle$ in Eq. (5.11) indicates that a particle, removed from the state l, is added to the state m. Thus, $i\omega = E_m - E_l - \mu$ means the corresponding excitation energy measured from μ, and the poles of $\tilde{G}_0(\mathbf{k}, i\omega)$ represents the energies of excitons.

15.6. Applications of a One-Body Time Green's Function

The one-body *time Green's function* can be defined in the time-coordinate space by the creation and annihilation operators $\Psi^{+}(\mathbf{r}, t)$, $\Psi(\mathbf{r}, t)$ in the Heisenberg representation as follows:

$$G(\mathbf{r}t; \mathbf{r}'t') = -i\langle T\Psi(\mathbf{r}, t)\Psi^{+}(\mathbf{r}', t')\rangle. \tag{6.1}$$

The choice of a factor $-i$ in this expression is in conformity with the differential equation of the Green's function G_0 for a free particle:

$$\left(i\frac{\partial}{\partial t} + \frac{\partial^2}{\partial \mathbf{r}^2}\right) G_0(\mathbf{r}, t) = \delta(\mathbf{r})\, \delta(t).$$

The Green's function thus defined has a dimension of $1/V$. If spins are involved, the field operators should be $\Psi_\sigma(\mathbf{r}, t)$ and $\Psi_\sigma^{*}(\mathbf{r}', t')$.

The Green's function can be used for evaluating the average of a one-body function B as follows (2):

$$\langle B \rangle = \pm i \int \lim_{\substack{\mathbf{r}' \to \mathbf{r} \\ t' \to t+}} B(\mathbf{r}, t) G(\mathbf{r}t; \mathbf{r}'t')\, d\mathbf{r}, \tag{6.2}$$

where the plus and the minus signs correspond respectively to bosons and fermions. For instance, the number and current densities can be expressed by

$$\begin{aligned} n(\mathbf{r}, t) &= \pm i \lim_{\substack{\mathbf{r}' \to \mathbf{r} \\ t' \to t+}} G(\mathbf{r}t; \mathbf{r}'t'), \\ \mathbf{j}(\mathbf{r}) &= \pm 2 \lim_{\substack{\mathbf{r}' \to \mathbf{r} \\ t' \to t+}} (\nabla_{\mathbf{r}} - \nabla_{\mathbf{r}'}) G(\mathbf{r}t; \mathbf{r}'t'). \end{aligned} \tag{6.3}$$

The factor 2 is due to our choice of units such that $2m = 1$. In other words, it is to be replaced by $1/m$ when ordinary units are used.

In homogeneous, infinite systems, the Green's functions depend only on the coordinate difference $\mathbf{r} - \mathbf{r}'$ and the time difference $t - t'$. For

$$G(\mathbf{r}' - \mathbf{r}; t' - t),$$

let us introduce

$$\begin{aligned} \tilde{G}(\mathbf{k}, t) &= \int (\exp -i\mathbf{k} \cdot \mathbf{r}) G(\mathbf{r}, t)\, d\mathbf{r}, \\ \tilde{G}(\mathbf{k}, \omega) &= \int_{-\infty}^{\infty} (\exp i\omega t) \tilde{G}(\mathbf{k}, t)\, dt. \end{aligned} \tag{6.4}$$

Note that the same function is used for the two transforms. However, no confusion will arise because we use the variable t for the first transform and ω for the second transform. The function $\tilde{G}(\mathbf{k}, t)$ is dimensionless and $\tilde{G}(\mathbf{k}, \omega)$ is of dimension t.

One can prove that the ground-state energy is given by

$$E_g = \frac{1}{4\pi i} \sum_{\mathbf{k}} \int d\omega\, (\omega + \omega_k) \tilde{G}(\mathbf{k}, \omega) \qquad (\omega_k = k^2). \tag{6.5}$$

The one-particle Green's function $\tilde{G}_0(\mathbf{k}, t)$ of a noninteracting fermion in the ground state is

$$\begin{aligned} \tilde{G}_0(\mathbf{k}, t) &= -\frac{i}{V} \langle 0 | T\left(\sum_{\mathbf{k}'} a_{\mathbf{k}'} a_{\mathbf{k}'}^{+} (\exp -i\omega_{k'} t)\right) \int [\exp i(\mathbf{k}' - \mathbf{k}) \cdot \mathbf{r}]\, d\mathbf{r} | 0 \rangle \\ &= \begin{cases} -i\langle 0 | a_{\mathbf{k}} a_{\mathbf{k}}^{+} | 0 \rangle (\exp -i\omega_k t), & t > 0, \\ i\langle 0 | a_{\mathbf{k}}^{+} a_{\mathbf{k}} | 0 \rangle (\exp -i\omega_k t), & t < 0, \end{cases} \\ &= \begin{cases} -i(1 - N_k) (\exp -i\omega_k t), & t > 0, \\ iN_k (\exp -i\omega_k t), & t < 0. \end{cases} \end{aligned} \tag{6.6}$$

We observe a discontinuous property

$$\tilde{G}_0(\mathbf{k}, +0) - \tilde{G}_0(\mathbf{k}, -0) = -i. \tag{6.7}$$

At 0°K, the occupation number of free fermions is given by

$$N_k = \begin{cases} 1, & |\mathbf{k}| < k_0, \\ 0, & |\mathbf{k}| > k_0. \end{cases}$$

Therefore, the Fourier transform defined by Eq. (6.4) becomes

$$\tilde{G}_0(\mathbf{k}, \omega) = -i\left[\theta(|\mathbf{k}| - k_0)\int_0^\infty e^{i(\omega-\omega_k)t}\,dt - \theta(k_0 - |\mathbf{k}|)\int_0^\infty e^{-i(\omega-\omega_k)t}\,dt\right]. \tag{6.8}$$

where

$$\theta(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0. \end{cases} \tag{6.9}$$

We can use the mathematical formula

$$\lim_{\varepsilon\to 0}\int_0^\infty e^{ist-\varepsilon t}\,dt = i\lim_{\varepsilon\to 0}\frac{1}{s+i\varepsilon} \tag{6.10}$$

in the integrals appearing in Eq. (6.8). We obtain

$$\tilde{G}_0(\mathbf{k}, \omega) = \lim_{\varepsilon\to 0}\left[\frac{\theta(|\mathbf{k}| - k_0)}{\omega - \omega_k + i\varepsilon} + \frac{\theta(k_0 - |\mathbf{k}|)}{\omega - \omega_k - i\varepsilon}\right] \tag{6.11}$$

$$= \lim_{\varepsilon\to 0}\left[\frac{1}{\omega - \omega_k + i\varepsilon\,\mathrm{sgn}(|\mathbf{k}| - k_0)}\right]. \tag{6.12}$$

In these expressions, ε is a convergence parameter which specifies the way to bypass the pole in the denominator of Eq. (6.12).

When an expression such as Eq. (6.12) is obtained, one can use an integral formula for $\varepsilon \to 0$

$$\int F(\omega)\frac{d\omega}{\omega + i\varepsilon} = P\int\frac{F(\omega)}{\omega}\,d\omega - i\pi F(0) \tag{6.13}$$

to evaluate $\tilde{G}_0(\mathbf{k}, t)$. Here, P means principal value. Note that the ω-integration for $t > 0$ can be performed by taking a contour integral along an infinite semicircle in the lower half-plane. For $t < 0$, the path must be in the upper half-plane. The results of these integrations agree with Eq. (6.6).

It should be remarked in this connection that if ε in the denominator of $\tilde{G}_0(\mathbf{k}, \omega)$ were a complex quantity, we would have obtained an expression such as

$$\tilde{G}_0(\mathbf{k}, t) \sim e^{-i\omega_k t}e^{-\lambda t}, \qquad t > 0. \tag{6.14}$$

This indicates that the corresponding quasi particle exists only for a limited interval of time. Here, $1/\lambda$ is the lifetime of the particle, which is determined by interactions.

Let us now examine the general Fourier transform $\tilde{G}(\mathbf{k}, \omega)$ by taking the matrix elements of Eq. (6.1) in an analogous way to the derivation of Eq. (5.11). The coordinate-space integration yields a delta function. The integration over t can be split into two intervals $(-\infty, 0)$ and $(0, \infty)$. We make use of

$$\int_0^\infty e^{i\omega t}\,dt = \pi\,\delta(\omega) + \frac{i}{\omega}. \tag{6.15}$$

For $t > 0$, we obtain

$$\int_0^\infty [\exp(i\omega t - i\mathbf{k}\cdot\mathbf{r})]G(\mathbf{r}, t)\,d\mathbf{r}\,dt = (2\pi)^3\sum_{l,m}[\exp\beta(\Omega + \mu N_l - E_l)] \times \delta(\mathbf{k} + \mathbf{k}_m{}^l)\langle l|\psi|m\rangle\langle m|\psi^+|l\rangle \times \left[\frac{1}{\omega + E_m{}^l} - i\pi\,\delta(\omega + E_m{}^l)\right],$$

where

$$\mathbf{k}_m{}^l = \mathbf{k}_l - \mathbf{k}_m, \qquad E_m{}^l = E_l - E_m + \mu. \tag{6.16}$$

A similar expression can be obtained for $0 > t > -\infty$. Combining the contributions from the two time intervals, we obtain

$$\tilde{G}(\mathbf{k}, \omega) = -(2\pi)^3\sum_{l,m}[\exp\beta(\Omega + \mu N_l - E_l)\langle l|\psi|m\rangle\langle m|\psi^+|l\rangle\,\delta(\mathbf{k} - \mathbf{k}_l{}^m) \times \left[\frac{1}{E_l{}^m - \omega}(1 \pm \exp -\beta E_l{}^m) + i\pi\,\delta(\omega - E_l{}^m)(1 \mp \exp -\beta E_l{}^m)\right]. \tag{6.17}$$

Therefore, if we write

$$\tilde{G}(\mathbf{k}, \omega) = \tilde{G}'(\mathbf{k}, \omega) + i\tilde{G}''(\mathbf{k}, \omega) \tag{6.18}$$

we have

$$\tilde{G}'(\mathbf{k}, \omega) = \begin{cases} \dfrac{P}{\pi}\displaystyle\int_{-\infty}^{\infty}\frac{\tilde{G}''(\mathbf{k}, x)}{x - \omega}\coth\left(\frac{\beta x}{2}\right)dx & \text{(fermions)}, \\ \dfrac{P}{\pi}\displaystyle\int_{-\infty}^{\infty}\frac{\tilde{G}''(\mathbf{k}, x)}{x - \omega}\tanh\left(\frac{\beta x}{2}\right)dx & \text{(bosons)}. \end{cases} \tag{6.19}$$

These two relations connect the real and imaginary parts. From Eq. (6.17), we find that $\tilde{G}''$ is always negative for bosons, but changes sign for fermions.

15.7. Response Functions

The Fourier transform $\chi_{BA}(\omega)$ of the response function $\phi_{BA}(t)$ given by Eq. (2.3) of Chapter 13 can be expressed by what is known as the *retarded Green's function*. To obtain such an expression, we integrate Eq. (2.3) of Chapter 13 by parts with the condition $\phi_{BA}(\infty) = 0$. We obtain

$$\chi_{BA}(\omega) = \frac{\phi_{BA}(0)}{i\omega} + \int_0^\infty \frac{e^{-i\omega t}}{i\omega} \frac{d\phi_{BA}}{dt}\, dt$$

$$= \int_0^\infty \frac{e^{-i\omega t} - 1}{i\omega} \frac{d\phi_{BA}}{dt}\, dt. \tag{7.1}$$

The time derivative $\dot{\phi}$ is given by Eq. (1.25) of Chapter 13. With $\hbar = 1$, it is

$$\frac{d\phi_{BA}}{dt} = -i \operatorname{tr} \rho_0[B(t), \dot{A}]. \tag{7.2}$$

For two operators A and B, the retarded Green's function is defined by

$$G^{\mathrm{R}}(t', t'') = -i\theta(t' - t'')\langle B(t')A(t'') - A(t'')B(t')\rangle, \tag{7.3}$$

where the function $\theta(x)$ is defined by Eq. (6.9) and

$$B(t') = (\exp i\mathscr{H}t')B(\exp -i\mathscr{H}t'), \quad \text{etc.} \tag{7.4}$$

The function $G^{\mathrm{R}}(t', t'')$ depends only on the difference $t' - t''$. Indeed,

$$\begin{aligned} G^{\mathrm{R}}(t', t'') &= -i\theta(t' - t'') \sum_{l,m} z^{N_l}[\exp \beta(\Omega - E_l)] \\ &\quad \times \{\langle l|B|m\rangle\langle m|A|l\rangle[\exp i(E_m{}^l - N_m{}^l\mu)(t' - t'')] \\ &\quad - \langle l|A|m\rangle\langle m|B|l\rangle[\exp i([E_l{}^m - N_l{}^m\mu)(t' - t'')]\} \\ &= -i\theta(\tau) \sum_{l,m} z^{N_l}[\exp \beta(\Omega - E_l)]\langle l|B|m\rangle\langle m|A|l\rangle \\ &\quad \times [\exp i(E_m{}^l - N_m{}^l\mu)\tau][1 - \exp \beta(E_m{}^l - N_m{}^l\mu)], \end{aligned} \tag{7.5}$$

where $\tau = t' - t''$, $E_m{}^l = E_l - E_m$, $N_m{}^l = N_l - N_m$, and where the second term of the second equation is obtained by interchanging the dummy indices l and m. One can also show that $G^{\mathrm{R}}(t', t'')$ satisfies a differential equation which is similar to Eq. (1.10).

The Fourier transform of $G^{\mathrm{R}}(t', t'') = G^{\mathrm{R}}(\tau)$ can be defined by

$$\tilde{G}^{\mathrm{R}}(\omega) = \lim_{\varepsilon \to 0} \int_0^\infty d\tau\, G(\tau) e^{i\omega\tau - \varepsilon\tau}, \tag{7.6}$$

where a parameter ε has been introduced to secure a convergence at the upper limit of integration. In terms of $\tilde{G}^{\mathrm{R}}(\omega)$, Eq. (7.1) becomes

$$\chi_{BA}(-\omega) = i \frac{\tilde{G}^{\mathrm{R}}(\omega) - \tilde{G}^{\mathrm{R}}(0)}{\omega}. \tag{7.7}$$

Therefore, the Fourier transform of the response function is expressed by that of the retarded Green's function:

$$\tilde{G}^{\mathrm{R}}(\omega) = \sum_{l,m} \frac{z^{N_l}[\exp \beta(\Omega - E_l)][1 - \exp \beta(E_m{}^l - N_m{}^l\mu)]}{\omega - (E_l{}^m - N_l{}^m\mu) + i\varepsilon} \langle l|B|m\rangle\langle m|A|l\rangle. \tag{7.8}$$

We shall now show that the Fourier transform of the retarded Green's function can be obtained from the temperature Green's function defined by

$$G(\beta', \beta'') = -\langle TB(\beta')A(\beta'')\rangle. \tag{7.9}$$

This is a generalization of the definition of the free-particle Green's function given by Eq. (4.20). Since the response functions are characterized by commutators instead of anticommutators, Wick's symbol T has the same effect as Dyson's chronological operator. Equation (7.9) is

$$G(\beta', \beta'') = \begin{cases} -\langle B(\beta')A(\beta'')\rangle, & \beta' > \beta'', \\ -\langle A(\beta'')B(\beta')\rangle, & \beta' < \beta''. \end{cases} \tag{7.10}$$

$G(\beta', \beta'')$ depends only on $\beta' - \beta'' = \alpha$. Its Fourier transform is from Eq. (2.11)

$$\begin{aligned} \tilde{G}(i\omega_n) &= \int_0^\beta G(\alpha) e^{i\omega_n\alpha}\, d\alpha \\ &= \sum_{l,m} \frac{z^{N_l}[\exp \beta(\Omega - E_l)][1 - \exp \beta(E_m{}^l - N_m{}^l\mu)]}{i\omega_n - (E_l{}^m - N_l{}^m\mu)} \\ &\quad \times \langle l|B|m\rangle\langle m|A|i\rangle. \end{aligned} \tag{7.11}$$

This expression corresponds to Eq. (5.11). Note that we are considering commutators (for bosons) instead of anticommutators (for fermions). Comparing Eq. (7.11) with Eq. (7.8), we find

$$\tilde{G}^{\mathrm{R}}(\omega) = \lim_{\varepsilon \to 0} \tilde{G}(\omega + i\varepsilon). \tag{7.12}$$

That is, the Fourier transform $\tilde{G}^{\mathrm{R}}(\omega)$ can be obtained from $\tilde{G}(\omega)$ by analytic continuation. Let us now consider two important examples.

15.7.1. Electrical Conductivity

The perturbation Hamiltonian is [see Eq. (5.1) of Chapter 13]

$$H_1 = -\mathbf{A} \cdot \mathbf{E}(t), \qquad \mathbf{A} = c \sum_i \mathbf{r}_s. \tag{7.13}$$

The response of the current density $\mathbf{B} = e\sum \dot{\mathbf{r}}_i$ to this perturbation can be expressed using Eq. (1.23) of Chapter 13. With the units $\hbar = 1$ and $2m = 1$, the time derivative of the response function is [see Eq. (1.25) of Chapter 13]

$$\frac{d\phi_{BA}}{dt} = -i \operatorname{tr} \rho_0[\mathbf{J}(t), \mathbf{J}]. \tag{7.14}$$

We define

$$\mathbf{G}^{\mathrm{R}}(\tau) = -i\theta(\tau)\langle[\mathbf{J}(\tau)\mathbf{J} - \mathbf{J}\mathbf{J}(\tau)]\rangle. \tag{7.15}$$

Then, the frequency dependent conductivity tensor is given by

$$\sigma(-\omega) = \frac{i}{\omega}[\tilde{\mathbf{G}}^{\mathrm{R}}(\omega) - \tilde{\mathbf{G}}^{\mathrm{R}}(0)] \qquad (V = 1 \text{ cm}^3). \tag{7.16}$$

Following the general argument of this section, the retarded Green's function can be obtained from Eq. (7.12). Since

$$\mathbf{J} = 2e\sum_{\mathbf{k}\sigma} \mathbf{k} a^{+}_{\mathbf{k}\sigma} a_{\mathbf{k}\sigma}, \tag{7.17}$$

the temperature Green's function is

$$\mathbf{G}(\alpha) = 4e^2 \sum_{\mathbf{k}\mathbf{k}'} \mathbf{k}\mathbf{k}' K(\mathbf{k}, \mathbf{k}'; \alpha), \tag{7.18}$$

where

$$K(\mathbf{k}, \mathbf{k}'; \alpha) = -\langle T a^{+}_{\mathbf{k}'}(\alpha) a_{\mathbf{k}'}(\alpha) a_{\mathbf{k}}{}^{+} a_{\mathbf{k}}\rangle_0. \tag{7.19}$$

This function has the same structure as the one that will be introduced in Eq. (8.7).

15.7.2. Dielectric Constant

The dielectric constant of an electron gas with a cloud of positive charges as background can be obtained by considering density fluctuations induced by a test charge (see Section 14.9). From Eq. (9.4) of Chapter 14, the interaction Hamiltonian is given by

$$\begin{aligned} H_1 &= -\sum_{\mathbf{k}} A(\mathbf{k}) e^{-i\omega t}, \\ A(\mathbf{k}) &= -\frac{4\pi e^2 I(k)}{k^2} n_{-\mathbf{k}} = -u(k) n_{-\mathbf{k}} I(k), \end{aligned} \tag{7.20}$$

where $u(k) = 4\pi e^2/k^2$. Since $F(t)$ of Eq. (1.3) of Chapter 13 is simply $e^{-i\omega t}$ in the present case, Eq. (1.19) of Chapter 13 is

$$\langle n_{\mathbf{k}}(t)\rangle = \int_0^{\infty} \phi_{BA}(\tau) e^{-i\omega\tau}\, d\tau\, e^{-i\omega t}. \tag{7.21}$$

In the absence of the test charge, there is no induced charge. Therefore, we have written $n_{\mathbf{k}}$ instead of $\Delta\langle n_{\mathbf{k}}\rangle$ in Eq. (7.21). According to Eq. (1.21b) of chapter 13, we have

$$\phi_{BA}(t) = -i \operatorname{tr} \rho_0[A(\mathbf{k}), n_{\mathbf{k}}(t)]. \tag{7.22}$$

Therefore, the Fourier transform of the density fluctuations is

$$\langle (n_{\mathbf{k}}(-\omega)\rangle = u(k) I(k) \tilde{G}^{\mathrm{R}}(\mathbf{k}, \omega), \tag{7.23}$$

where $\tilde{G}^{\mathrm{R}}(\omega)$ is the Fourier transform of the retarded Green's function:

$$G^{\mathrm{R}}(\tau) = -i\theta(\tau) \operatorname{tr} \rho_0[n_{\mathbf{k}}(\tau), n_{-\mathbf{k}}]. \tag{7.24}$$

On the other hand, Maxwell's equations give [see Eq. (9.10) of Chapter 14]

$$\begin{aligned} i\mathbf{k} \cdot \tilde{\mathbf{D}}(\mathbf{k}, \omega) &= 4\pi e I(k), \\ i\mathbf{k} \cdot \tilde{\mathbf{E}}(\mathbf{k}, \omega) &= 4\pi e I(k) + 4\pi e \langle n_{\mathbf{k}}(\omega)\rangle V^{-1}, \end{aligned} \tag{7.25}$$

where $\tilde{\mathbf{D}}$ and $\tilde{\mathbf{E}}$ are the Fourier transforms of the displacement and electric field vectors. The dielectric function $\varepsilon(\mathbf{k}, \omega)$ is defined by

$$\tilde{\mathbf{D}}(\mathbf{k}, \omega) = \varepsilon(\mathbf{k}, \omega)\tilde{\mathbf{E}}(\mathbf{k}, \omega). \tag{7.26}$$

Considering longitudinal modes, we find

$$\begin{aligned} [\varepsilon(\mathbf{k}, \omega)]^{-1} &= 1 + \frac{\langle n_{\mathbf{k}}(\omega)\rangle}{VI(k)} \\ &= 1 + \frac{u(k)\tilde{G}^{\mathrm{R}}(\mathbf{k}, -\omega)}{V}. \end{aligned} \tag{7.27}$$

As before, the Fourier transform of the retarded Green's function $G^{\mathrm{R}}(\tau)$ for the dielectric function is obtained from the temperature Green's function defined by

$$G(\mathbf{k}, \alpha) = -\langle T n_{\mathbf{k}}(\alpha) n_{-\mathbf{k}}(0)\rangle, \tag{7.28}$$

where we have used a simplified notation $G(\mathbf{k}, \alpha)$ instead of $G(\mathbf{k}; \beta', \beta'')$ for $\alpha = \beta' - \beta''$. Equation (7.28) can be approximately evaluated as follows:

$$\begin{aligned} G(\mathbf{k}, \alpha) &\approx -\sum_{\mathbf{q},\sigma} \langle a^{+}_{\mathbf{q}-\mathbf{k},\sigma}(\alpha) a_{\mathbf{q}-\mathbf{k},\sigma}\rangle_0 \langle a_{\mathbf{q},\sigma}(\alpha) a^{+}_{\mathbf{q},\sigma}\rangle_0 \\ &= 2\sum_{\mathbf{q}} G_0(\mathbf{q} - \mathbf{k}; 0, \alpha) G_0(\mathbf{q}; \alpha, 0), \end{aligned} \tag{7.29}$$

where G_0 is given by Eq. (3.17). The Fourier transform of Eq. (7.29) is

$$\begin{aligned} \tilde{G}(\mathbf{k}, i\omega_l) &= \int_0^{\beta} G(\mathbf{k}; \alpha, 0) e^{i\omega_l \alpha}\, d\alpha \\ &= \frac{2}{\beta}\sum_{\mathbf{q},m} [(i\omega_l - \varepsilon^0_{\mathbf{q}-\mathbf{k}} + \mu)(i\omega_{l+m} - \varepsilon_q{}^0 + \mu)]^{-1} \\ &= -V\lambda_l(k), \end{aligned} \tag{7.30}$$

because of Eq. (4.25). Note here that $\omega_l = 2\pi l/\beta$. Finally, we obtain

$$[\varepsilon(\mathbf{k}, \omega_l)]^{-1} = 1 - u(k)\lambda_l(k). \tag{7.31}$$

The dielectric function is determined by $\lambda_l(k)$, which has been given by Eq. (8.9) of Chapter 10 for degenerate electrons.

15.8. Equations of Motion

Using the equations of motion for the Heisenberg operator, we can derive the equations of motion for the Green's functions. We start with the equation of motion for the Heisenberg operator Ψ:

$$\begin{aligned} i\frac{\partial}{\partial t}\Psi(\mathbf{r}, t) + (\nabla^2 + \mu)\Psi(\mathbf{r}, t) &= [\Psi(\mathbf{r}, t), H_1], \\ \Psi(\mathbf{r}, t) &= e^{it\mathscr{H}}\psi(\mathbf{r})e^{-it\mathscr{H}}, \\ \Psi^+(\mathbf{r}, t) &= e^{it\mathscr{H}}\psi^+(\mathbf{r})e^{-it\mathscr{H}}, \end{aligned} \tag{8.1}$$

where $\mathscr{H} = H - \mu N$.

In terms of a ground-state average $\langle \cdots \rangle$, we define

$$\begin{aligned} \langle T\Psi(\mathbf{r}, t)\Psi^+(\mathbf{r}', t')\rangle &= \theta(t - t')\langle \Psi(\mathbf{r}, t)\Psi^+(\mathbf{r}', t')\rangle \\ &\quad - \theta(t' - t)\langle \Psi^+(\mathbf{r}', t')\Psi(\mathbf{r}, t)\rangle, \end{aligned} \tag{8.2}$$

where $\theta(t) = 1$ for $t > 0$ and $\theta(t) = 0$ for $t < 0$. Differentiating Eq. (8.2) with respect to time, we find

$$\begin{aligned} i\frac{\partial}{\partial t}G(\mathbf{r}t, \mathbf{r}'t') &= \theta(t - t')\left\langle \frac{\partial}{\partial t}\Psi(\mathbf{r}, t)\Psi^+(\mathbf{r}', t')\right\rangle \\ &\quad - \theta(t' - t)\left\langle \Psi^+(\mathbf{r}', t')\frac{\partial\Psi(\mathbf{r}, t)}{\partial t}\right\rangle \\ &\quad + \delta(t - t')\langle[\Psi(\mathbf{r}, t)\Psi^+(\mathbf{r}', t') + \Psi^+(\mathbf{r}', t')\Psi(\mathbf{r}, t)]\rangle. \end{aligned} \tag{8.3}$$

Thus,

$$\left(i\frac{\partial}{\partial t} + \nabla^2 + \mu\right)G(\mathbf{r}t, \mathbf{r}'t') = \delta(\mathbf{r} - \mathbf{r}')\,\delta(t - t') - i\langle T[\Psi(\mathbf{r}, t), H_1]\Psi^+(\mathbf{r}', t')\rangle. \tag{8.4}$$

If the interaction H_1 is independent of the spin variable and is given by

$$H_1 = \frac{1}{2}\int d\mathbf{r}\, d\mathbf{r}'\, \psi^+(\mathbf{r}')\psi^+(\mathbf{r})\phi(|\mathbf{r}' - \mathbf{r}|)\psi(\mathbf{r})\psi(\mathbf{r}'), \tag{8.5}$$

then

$$[\Psi, H_1] = \int d\mathbf{r}'\, \phi(|\mathbf{r}' - \mathbf{r}|)\psi^+(\mathbf{r}')\psi(\mathbf{r}')\psi(\mathbf{r}). \tag{8.6}$$

The Heisenberg representation can be used for these equations.

A two-particle Green's function is defined by a ground-state average:

$$K(1234) = \langle T(\Psi(1)\Psi(2)\Psi^+(3)\Psi^+(4))\rangle, \tag{8.7}$$

where 1, 2, etc. represent $(\mathbf{r}_1, t_1)$, $(\mathbf{r}_2, t_2)$, etc. Equation (8.4) becomes

$$\begin{aligned} \left(i\frac{\partial}{\partial t} + \nabla^2 + \mu\right)G(\mathbf{r}t \cdot \mathbf{r}'t') &= \delta(t - t')\,\delta(\mathbf{r} - \mathbf{r}') \\ &\quad - i\int d\mathbf{r}''\, \phi(|\mathbf{r} - \mathbf{r}''|)K(\mathbf{r}''t, \mathbf{r}t, \mathbf{r}'t', \mathbf{r}''t+). \end{aligned} \tag{8.8}$$

Thus, the Green's functions form a hierarchy of equations: In order to determine the one-particle Green's function, one needs the two-particle Green's function, and so on.

In Eq. (8.8), let us take the limit $\mathbf{r} \to \mathbf{r}'$ and $t' \to t+$ and integrate over $\mathbf{r}$. We have

$$\begin{aligned} \langle H_1\rangle &= \frac{1}{2}\int d\mathbf{r}\, d\mathbf{r}''\, \phi(|\mathbf{r} - \mathbf{r}''|)K(\mathbf{r}''t, \mathbf{r}t, r't', \mathbf{r}''t^+) \\ &= -\frac{i}{2}\int d\mathbf{r} \lim_{\substack{\mathbf{r}'\to\mathbf{r} \\ t'\to t+}} \left(i\frac{\partial}{\partial t} + \nabla^2 + \mu\right)[G(\mathbf{r} - \mathbf{r}', t - t') - G_0(\mathbf{r} - \mathbf{r}', t - t')]. \end{aligned} \tag{8.9}$$

15.9.* Superconductivity

Superconductivity was discovered in 1911 by Kamerlingh Onnes in Leiden. He observed that the residual electrical resistance of mercury dropped to zero at 4.15°K. To reduce the resistance caused by impurities, he used highly purified mercury, and therefore he was not surprised to find near-zero resistance. However, it was a surprise to observe that the decrease took place within 0.01°K of the finite temperature and that the resistance did not increase much even when impurities were added. Thus, after a few years of additional study, he concluded that at the temperature mercury passed into a new state, called the superconductive state.

Since then, superconductivity has been discovered in many metals, compounds, and alloys. The temperature at which the transition takes place is usually of the order of 1°K. One of the highest measured transition

temperatures is 20.7°K of $Nb_{0.8}(Al_{0.75}Ge_{0.25})_{0.2}$. The best material known at present for a high-field superconducting magnet is Nb_3Sn, which has a critical field of about 220 kG at 4.2°K. A critical field of 320 kG has been reported for Nb_3Al at 4.2°K.

Perhaps the most dramatic experiment is on the persistent current induced in a superconducting ring. In an experiment at M.I.T., a current was found to circulate without appreciable diminution for more than a year. Thus, we conclude that the electrical resistance is essentially zero; it remains so up to megacycle frequencies.

We shall list several important experimental facts which were found prior to development of the Bardeen–Cooper–Schrieffer (BCS) theory (*8*).

(*i*) *Magnetic Field effect (Kamerlingh–Onnes)*

Superconductivity may be destroyed by a magnetic field. The minimum magnetic field required to destroy superconductivity obeys the formula

$$\frac{H_c}{H_c(0)} = 1 - \left(\frac{T}{T_c}\right)^2. \tag{9.1}$$

(*ii*) *Meissner–Ochenfeld effect*

Superconductors are perfectly diamagnetic. In other words, the magnetic induction B inside a superconductor is zero. The penetration depth of a magnetic field is of the order of $10^{-5} \sim 10^{-6}$ cm.

(*iii*) *Order of transition*

The phase transition to the superconducting state in the absence of a magnetic field is in second order. The electronic specific heat jumps to several times the normal value at the transition. The presence of a magnetic field makes the transition first order, i.e., a latent heat is involved.

(*iv*) *Isotope effect*

In 1950, E. Maxwell and others found the *isotope effect*: The transition temperature is dependent on the isotopic mass M as follows

$$T_c \propto M^{-1/2}. \tag{9.2}$$

The isotope effect gave a clue that superconductivity was caused by electron–electron interactions via phonons. Fröhlich was the first to give a theoretical treatment for electrons in metals with lattice vibrations and to conclude that the interaction led to instability of the Fermi surface. Thus, the filled Fermi sea is no longer the lowest state. However, Fröhlich's approach was based essentially on an independent-electron model and was found inadequate to explain the cooperative features of the transition. In 1956, Cooper showed that two electrons, attracting each other by exchanging phonons, could form a bound state. This result was the springboard for the famous BCS theory, published the next year (*8*).

Since the BCS theory, many other important theories have been published. It is definitely beyond the scope of this book to discuss the details of these theories, but in concluding this chapter we shall outline the Green's function method due to Gorkov, published in 1958 (*9*). In this theory, the electron–electron interactions via phonons are simply replaced by constant attractive potentials as in the BCS theory. Thus, the effective Hamiltonian can be written from Eq. (7.9) of Chapter 14 as (10)

$$H = \int d\mathbf{r}\, \Psi^+(\mathbf{r}, t) p^2 \Psi(\mathbf{r}, t) - \tfrac{1}{2}\phi \int d\mathbf{r}\, d\mathbf{r}'\, \Psi^+(\mathbf{r}, t)\Psi^+(\mathbf{r}', t)\, \delta(\mathbf{r} - \mathbf{r}')\Psi(\mathbf{r}', t)\Psi(\mathbf{r}, t). \tag{9.3}$$

The potential ϕ is zero except in a narrow energy shell about the Fermi surface.

The equation of motion (8.8) can be simplified. The potential can be taken outside of the integrand. Now, $\phi(r)$ becomes $-\phi = \text{const}$. In the Hartree–Fock approximation, the two-body Green's function can be replaced by products of one-body Green's functions. In addition, we can include a special term which represents the creation and annihilation of bound electron pairs. Furthermore, the operators $\Psi\Psi$ and $\Psi^+\Psi^+$ for these processes can be treated as c-numbers if the number of bound pairs is large. In other words, we use

$$T\langle\Psi(1)\Psi(2)\Psi^+(3)\Psi^+(4)\rangle = -G(1, 3)G(2, 4) + G(1, 4)G(2, 3) + F(1, 2)F^+(3, 4) \tag{9.4}$$

in Eq. (8.8). We obtain

$$\left(i\frac{\partial}{\partial t} + \nabla^2 + \mu\right)G(x - x') + i\phi F(+0)F^+(x - x') = \delta(x - x'), \tag{9.5}$$

where x represents a set of coordinates $(\mathbf{r}, t)$.

In a way similar to the derivation of Eq. (8.8) for $G(x)$, one can obtain the equation of motion for $\Psi(\mathbf{r}, t)\Psi^+(\mathbf{r}', t')$, which is $F^+(\mathbf{r} - \mathbf{r}')$. The result is

$$\left(-i\frac{\partial}{\partial t} + \nabla^2 + \mu\right)F^+(x - x') + i\phi F^+(+0)G(x - x') = 0. \tag{9.6}$$

Expressing $F(x)$ and $G(x)$ in Fourier transforms, we find

$$\begin{aligned}(\omega - k^2 + \mu)\tilde{G}(\mathbf{k}, \omega) + i\phi F(+0)\tilde{F}^+(\mathbf{k}, \omega) &= 1,\\ (\omega + k^2 - \mu)\tilde{F}^+(\mathbf{k}, \omega) - i\phi F^+(+0)\tilde{G}(\mathbf{k}, \omega) &= 0.\end{aligned} \tag{9.7}$$

Thus, we obtain

$$\tilde{G}(\mathbf{k}, \omega) = \frac{\omega + \bar{\varepsilon}_k}{\omega^2 - (\bar{\varepsilon}_k^2 + \Delta_0^2)},$$
$$\tilde{F}^+(\mathbf{k}, \omega) = \frac{i\phi F^+(+0)}{\omega^2 - (\bar{\varepsilon}_k^2 + \Delta_0^2)}, \tag{9.8}$$

where $\bar{\varepsilon}_k = k^2 - \mu$ and $\Delta_0^2 = F(+0)F^+(+0)\phi^2$.

We obtain a similar set of equations for finite temperatures by replacing t by $-i\alpha$ in Eqs. (9.5) and (9.6) and using a Fourier series such as appeared in Eq. (2.8) instead of a Fourier integral over the time variable. The results are given by replacing ω of Eqs. (9.8) by $i\omega_n = (2n+1)\pi i/\beta$ and $i\phi$ by $-\phi$:

$$\tilde{G}(\mathbf{k}, i\omega_n) = -\frac{i\omega_n + \bar{\varepsilon}_k}{\omega_n^2 + (\bar{\varepsilon}_k^2 + \Delta^2)},$$
$$\tilde{F}^+(\mathbf{k}, i\omega_n) = \frac{\phi F^+(+0)}{\omega_n^2 + (\bar{\varepsilon}_k^2 + \Delta^2)}. \tag{9.9}$$

The one-particle Green's function $\tilde{G}(\mathbf{k}, i\omega_n)$ can be reduced to an expression which is similar to that for an ideal Bose gas given by Eq. (3.16). (This equation was derived for fermions, but it holds for bosons also if n is even.) The denominator of $\tilde{G}(\mathbf{k}, i\omega_n)$ has a positive imaginary pole at

$$\omega_n = i(\bar{\varepsilon}_k^2 + \Delta^2)^{1/2}. \tag{9.10}$$

$(\bar{\varepsilon}_k^2 + \Delta^2)^{1/2}$ may be considered as the quasi-particle energy. This energy is characterized by Δ, which represents an energy gap.

The energy gap parameter Δ is a function of the temperature. When Δ is zero, a phase transition similar to the Bose condensation takes place. To determine Δ, we use the second of Eqs. (9.9) or Eq. (9.8). From the Fourier transform relation

$$F^+(+0) = \frac{1}{(2\pi)^3\beta}\sum_n \int d\mathbf{k}\, F^+(\mathbf{k}, i\omega_n), \tag{9.11}$$

we obtain the self-consistent relation

$$1 = \frac{\phi k T}{(2\pi)^3}\sum_n \int \frac{d\mathbf{k}}{\omega_n^2 + (\bar{\varepsilon}_k^2 + \Delta^2)}, \tag{9.12}$$

which determines the gap parameter.

The momentum-space integration of Eq. (9.12) can be limited to energies in a layer of thickness of $2\omega_D$ about the Fermi surface since only electrons in such a layer participate in the interaction. The transition temperature is found to be proportional to the thickness and is given by

$$T_c = \frac{\gamma}{\pi k}(2\omega_D)\exp\frac{-2\pi^2}{\phi m p_0}, \tag{9.13}$$

where $\log \gamma = 0.577$ is Euler's constant, p_0 is the Fermi momentum, and we have made explicit the electron mass m. In terms of T_c, the gap parameter is found to be

$$\Delta = \Delta_0 - (2\pi k T \Delta_0)^{1/2}\left(1 - \frac{kT}{8\Delta_0}\right)e^{-\Delta_0/kT} \tag{9.14}$$

for low temperatures, and

$$\Delta = 3.06 T_c\left(1 - \frac{T}{T_c}\right)^{1/2} \tag{9.15}$$

near the transition temperature. Here, $\Delta_0 = \phi|F^+(+0)|$ represents the energy gap at absolute zero.

Correspondingly, the thermodynamic properties of the system change at $T = T_c$. For instance, it can be shown that the heat capacity at very low temperatures varies in accordance with

$$C(T) = C_n(T_c)\frac{3}{\gamma}\left(\frac{2}{\pi}\right)^{1/2}\left(\frac{\Delta_0}{kT}\right)^{3/2}e^{-\Delta_0/kT}, \tag{9.16}$$

while in the vicinity of the transition temperature, the variation is

$$C(T) = C_n(T_c)\left[2.43 + 3.77\left(\frac{T}{T_c} - 1\right)\right]. \tag{9.17}$$

Here,

$$C_n(T) = \tfrac{1}{3} m p_0 k^2 T \tag{9.18}$$

is the heat capacity per unit volume of the normal metal; Eq. (9.18) has been derived in Section 4.6. Thus, the specific heat at constant volume jumps at the transition temperature in agreement with experiment. Moreover, one can show that the critical magnetic field varies with the temperature as

$$H_c(T) = H_c(0)\left(1 - \frac{\gamma^2}{3}\frac{T^2}{T_c^2}\right) \tag{9.19}$$

for $T \ll T_c$, and

$$H_c(T) = 1.73\, H_c(0)\left(1 - \frac{T}{T_c}\right). \tag{9.20}$$

for $T \leq T_c$. These results are in satisfactory agreement with experiments, as can be seen from the empirical formula (9.1) (11).

REFERENCES

1. V. L. Bonch-Bruevich and S. V. Tyablikov, "The Green Function Method in Statistical Mechanics" (translated by D. ter Haar). North-Holland, Amsterdam, 1962.
2. A. A. Abrikosov, L. P. Gorkov, and I. E. Dzyaloshinski, "Quantum Field Theory in Statistical Physics" (translated by R. A. Silverman). Prentice Hall, Englewood Cliffs, New Jersey, 1963.
3. J. Schwinger, *Proc. Nat. Acad. Sci. U.S.A.* **37**, 452 (1951).
4. P. C. Martin and J. Schwinger, *Phys. Rev.* **118**, 1417 (1959).
5. W. Kohn and J. M. Luttinger, *Phys. Rev.* **124**, 670 (1960).
6. D. N. Zubarev, *Soviet Phys.—Uspekhi* **3**, 320 (1960).
7. W. E. Parry and R. E. Turner, *Rep. Progr. Phys.* **27**, 23 (1964).
8. J. Bardeen, L. N. Cooper, and J. R. Schrieffer, *Phys. Rev.* **108**, 1175 (1957).
9. L. P. Gorkov, *Soviet Physics JETP* **34**, 505 (1958).
10. J. R. Schrieffer, "Theory of Superconductivity." Benjamin, New York, 1964.
11. R. D. Parks, ed., "Superconductivity." Dekker, New York, 1969.

Subject Index

A

A priori probability assumption, 20, 39, 56
Absorption coefficient, 70
Activity, 61, 62, 124, 126, 170
  absolute, 61, 62, 289
Adiabatic expansion, 119, 120
Adiabatic process, 96
Alloy, 266, 271
Annihilation operator, 383
Antiferromagnetism, 258
Autocorrelation function, 199
Average speed, 9, 99
Average velocity, 154

B

BBGKY hierarchy, 155, 197
BCS theory, 428, 429
Bernoulli equation, 84
Bethe Ansatz, 335
Binary mixture, 148, 259
Blackbody radiation, 71
Bloch's equation, 280, 282, 362, 413
Bogoliubov transformation, 389
Bohr magneton, 103
Bohr radius, 325, 383
Boltzmann constant, 33
Boltzmann equation, 3, 6, 156, 158, 338
Boltzmann factor, 51, 285
Boltzmann H-function, 6
Boltzmann relation, 51
Boltzmann statistics, 49, 288
Born–Green and Kirkwood equation, 177
Bose–Einstein condensation, 84, 86, 106, 388, 421, 423
Bose–Einstein distribution, 82
Bose–Einstein gas, 80, 87, 98, 249, 313, 328, 347, 430
Bose–Einstein statistics, 80, 105, 248, 288, 339, 398, 400, 402
Bose function, 88, 342
Boson, 80, 82, 282, 303, 314, 319, 327, 358
Bounds of free energy, 147
Boyle temperature, 116
Bragg–Williams theory, 263, 268
Brownian motion, 159, 182, 201
Brownian particle, 193, 195, 196, 198, 212
Bulk viscosity, 160

C

Canonical distribution, 33, 367
Canonical ensemble, 52, 53, 280
Cauchy–Schwartz inequality, 151
Chain diagram, 172
  approximation, 317

Chapman–Kolmogorov equation, 206
Charged boson, 327
Chemical potential, 61, 261
  of solvent, 150
Clausius–Mossotti formula, 143
Clausius' virial theorem, 18, 164
Closed polygon, 219
Cluster expansion, 113, 137, 148, 169, 209, 236, 293ff, 311, 313, 375
Cluster integral, 120ff, 122
Coarse grained distribution function, 38, 40
Coarse graining process, 37, 39
Collision cylinder, 4, 5
Collision hypothesis, 4
Collision term, 155, 339
Compressibility, 168, 256, 268, 270
Condensation, 84, 86, 106, 136
Conductivity
  electrical, 69, 271, 352ff, 369, 373ff, 423
  thermal, 15, 16, 100, 340, 345, 346, 352, 355
Configurational partition function, 114, 121ff, 125, 253
Connected graph, 121
Contraction, 409
Convolution approximation, 174
Cooperative phenomena, 263
Correlation energy, 304ff, 392
Correlation function, 160, 250
  density–density, 169, 204, 266
  space–time, 203
  spin, 230
  time, 373, 375
Creation operator, 383
Critical exponent, 256ff
Critical temperature, 139, 253ff, 428, 431
  divergence, 255, 256
  dip, 271
  slowing down, 271
  opalescence, 166
Cross, 224
Cross section, 5, 203, 339, 344, 351, 353
Cumulant theorem, 133, *see also* Thiele's semi-invariant,

D

Debye characteristic temperature, 78, 79, 95, 243
Debye function, 78, 79
Debye–Hückel distribution function, 136, 321
  correction to, 324
Debye–Hückel screening constant, 135
Debye model, 78, 239
Degenerate gas, 84, 343, 347, 388
de Haas–van Alphen effect, 101
Density fluctuation, 64, 167, 169, 266, 399, 424
Density matrix, 278ff
Density of states, 83
Detailed balance condition, 7, 44, 156, 352
Dielectric constant, 204, 424
  of plasma, 398, 426
Dielectric function, 399, 425, 426
Diffusion, 16, 186, 212, 271, 272, 370
  equation, 185, 211, 212
  in momentum space, 159, 200
Dipole interaction, 141
Distribution function, 21, 33, 153ff, 162, 169, 310ff
  canonical, 33, 367
  coarse-grained, 38
  Debye–Hückel, 136, 321
  fine grained, 38, 40
  frequency, 77, 240ff
  pair, *see* Pair distribution function
  phase space, 21, 280
  singlet, 154, 157, 376
  time smoothed, 198
  triplet, 316
  Wigner, 280
DNA melting, 217
Doppler broadening, 11
Dual lattice, 220
Duality principle, 221
Dulong–Petit's law, 79
Dumbbell molecule, 116, 117
Dynamic form factor, 397
Dynamic reversibility, 17
Dyson's time ordering parameter, 406

E

Eigenvalues of propagator, 298, 299ff
Einstein characteristic temperature, 76
Einstein function, 76
Einstein model, 75
Einstein relation, 194, 370
Einstein viscosity formula, 190, 192

Elasticity of rubber, 65
Electrical conductivity, 69, 271, 352ff, 369, 373, 423
Electron gas, 134ff, 320, 391
Electronic heat capacity, 94
  specific heat, 94
Electron interaction via phonon, 395
Electron phonon interaction, 393
Energy density, 73, 74
Energy gap, 430
Ensemble, 52
  average, 24, 26, 53, 56
  canonical, 24, 26, 53, 56
  grand canonical, 57
  microcanonical, 22
Enthalpy, 120
Entropy, 55, 264
  increase, 36
  statistical definition of, 33, 39
Equation of state
  Bernoulli's, 84
  of ideal gas, 8
  of imperfect gas, 114, 122ff, 125, 127
  of quantum gases, 74, 83, 84, 86, 92, 94, 292, 334
  van der Waals, 115
  virial, 164
Equipartition law, 10, 201, 239
Ergodic hypothesis, 24
Ergodic theory, 24, 25, 200
Exchange energy, 305, 390
Extensive variable, 56

F

f-bond, 121
f-sum rule, 397, 400
Fermi (Dirac) distribution, 82, 97, 302, 325, 339, 345, 350, 352, 356, 385, 415
Fermi (Dirac) statistics, 80, 105, 339, 401
Fermi energy, 305
Fermi function, 342
Fermi gas, 80, 86, 91, 99, 332, 343
Fermi liquid, 91, 96, 344
Fermi momentum, 92, 102, 302
Fermi sphere, 92
Fermion, 80, 82, 282, 302, 319, 325, 344, 415, 421, 423
Ferrimagnetism, 214
Ferromagnetism, 215, 248, 257, 264
Fine grained distribution function, 38
First law of thermodynamics, 35
First order phase transition, 252, 428
Fluctuating force, 192, 197, 199
Fluctuation, 63, 64, 149
  spin, 344
Fokker–Planck's equation, 198, 205ff
Fourth virial coefficient, 178
Free energy
  Helmholtz's, 35, 36, 53, 55, 62, 90, 108, 264
  Gibbs', 35, 61
  statistical definition of, 35
Free path, 12
Frenkel defect, 68
Frequency distribution function, 240
Frequency spectrum, 239ff
Friction constant (frictional coefficient), 159, 192, 195, 370
Fugacity, 61

G

Gamma space, 37
Gibbs ensemble, 53
Gibbs free energy, 61
Gibbs *H*-theorem, 31
Gibbs inequality, 37, 45, 144
Gibbs paradox, 55
Grand ensemble, 57, 123, 311, 384, 406
Grand partition function, 58, 59, 81, 123, 124, 142, 148, 164, 233, 293, 299, 412, 415
Grand potential, 60, 146, 406
Green's function, 211, 286, 361, 405ff
  differential equation for, 195, 407, 426
  lattice, 245
  retarded, 422ff
  temperature, 425ff, 423
  time, 418ff
Ground state energy, 249, 304, 306, 330, 334, 392, 419
Grüneisen–Bloch formula, 354

H

H-theorem, 6, 30–32, 40, 42, 43
  Boltzmann's, 6, 17, 30
  Gibbs', 31

Hard-sphere gas, 178, 217
  Bose, 329
  Fermi, 332
Harmonic oscillator, 72, 75, 284, 387, 400
Heat capacity, *see* Specific heat
Heisenberg model, 215, 247ff
Helix-coil transition, 217
Helmholtz free energy, 35
Henry's constant, 149
Henry's law, 125
Helium-three
  mixed with ^{4}He, 261
  refrigeration, 261
HNC approximation, 174
Hollow space radiation, 71

I

Ideal gas, 8, 36, 55, 56, 62, 64, 81
  mixture, 63, 125, 150
  solution, 150
Induced dipole, 140
Interaction representation, 413
Internal energy, 34, 52, 53, 75, 87, 106, 107, 163, 264
  of radiation, 74
  statistical definition of, 34, 53
  surface, 164
Intrinsic viscosity, 188
Invariant part of phase space, 25
Inverse collision, 6
Irreducible integral, 127, 128
Irreversibility, 39, 43, 156, 197, 198, 200, 360ff, 366, 376ff
Irreversible process, 13, *see also* Transport phenomena
Ising lattice, 214ff
Isothermal compressibility, 168, 256, 268, 270
Isotope effect, 428

J

Joule–Thomson effect, 119

K

Kinetic theory, 3ff, 18
Kirchhoff's law, 71
Kirkwood–Boggs equation, 180
Kramers–Kronig relations, 195, 365

L

Langevin equation, 159, 192, 194, 201
Lattice defect, 68, 247
Lattice dynamics, 75
Lattice gas, 232ff, 249, 263ff
Lattice vibration, 75, 215, 239ff, 387
Lattice walk, 182, 186
Law of corresponding states, 79, 116, 253
Lehmann representation, 416
Lennard-Jones potential, 115, 176
Liouville equation, 22, 153, 155, 162, 196
Liouville operator, 196, 361
Liouville theorem, 18, 22, 26, 361
Liquid solution, 148, 259ff, 261
Logarithmic divergence, 229, 255, 358
Long range order, 263
Lorentz group, 403
Loschmidt's number, 12

M

Magnon, 248, 344
Markoff chain, 206
Markoff method, 183, 185
Markoff process, 206
Master equation, 43, 44, 377
Maxwell–Boltzmann distribution, 7, 12ff, 192
Maxwell's law of equal areas, 253
Mean ergodic theorem, 25
Mean free path, 12, 98, 99
Meissner–Ochsenfeld effect, 428
Metrically indecomposable set, 25, 26
Microcanonical distribution, 20
Microcanonical ensemble, 22, 27
Mixed state, 279
Mixture, 148, 259ff
Mobility, 194
Møller operator, 308
Molecular chaos, 4, 377
Most probable velocity, 10

N

Nearest neighbor statistics, 215
Néel temperature, 258, 259
Nernst relation, 213

Neutron scattering, 89, 169, 186, 202ff, 242, 261
Nodal expansion, 172
Nodal point, 172
Normal process, 354
Number operator, 383
Number of states, 73, 82, 83, 101
Nyquist relation, 380

O

One-dimensional anharmonic oscillator, 70
One-dimensional gas, 217
One-half power law, 254
One-third power law, 254, 260
Onsager relations, 366
Order–disorder phenomena, 260, 263
Order of transition, 252
Ornstein–Zernike equation, 174
Orthogonal group, 403
Oseen formula, 189
Osmotic pressure, 148

P

Pair distribution function, 160, 163ff, 196, 203, 311, 312, 317ff
  for classical electron gas, 320ff, 334
  for fermions with spins, 335, 336
  for hard-sphere gas, 328, 332
  for ideal quantum gas, 270, 314, 332, 335
Parastatistics, 105, 402
Partition function, 51ff, 72, 81, 108, 113, 208, 215ff, 281, 285, 303
Pauli equation, 43, 377
Peierls inequality, 147
Pfaffian, 224
Phase space distribution, 21, 280
Phase point, 21, 27
Phase space, 21, 38
Phase transition
  first order, 139, 252, 428
  second order, 252, 428
Phase volume, 56
Phonon excitation, 329, 335, 348ff, 388
Phonon–phonon scattering, 349
Phonon–roton scattering, 351
Phonon spectrum, 89, 388ff
Plasma, 161, 398
Plasma parameter, 321
Poisson–Boltzmann equation, 151
Polarizability, 141
Polarizable gas, 140
Polarization, 394
Polaron, 393
Polyatomic gas, 16
Polymer solution, 150
Potential of average force, 177
Pressure broadening, 11
Propagator, 277, 286ff, 310ff
Pseudopotential, 204, 329
Pure state, 279
Py approximation, 174

Q

Quasi-ergodic hypothesis, 24, 25
Quasi-ergodicity, 30
Quasi-particle, 96, 262, 344, 420, 430

R

Radial distribution function, 160, 165, 166
Random walk, 186, 193, 200, 220
Rayleigh scattering, 108, 350
Rectilinear relation, 235, 254
Recurrence theorem, 17
Reduced distribution function, 42, 154, 157, 163
Regular solution, 150
Relativistic gas, 107
Relaxation time, 350, 353, 356, 360
Response function, 195, 362ff, 422ff
  dissipative, 364
  symmetry properties of, 366
Retarded Green's function, 422ff
Ring diagram approximation, 134, 295, 297
Roton, 89, 348
Roton-roton scattering, 351
Rubber elasticity, 65, 185
Rydberg unit, 305, 383

S

Schottky defect, 68
Second law of thermodynamics, 35, 53
Second order phase transition, 252, 428
Second quantization, 147, 382ff

Second virial coefficient, 147, 382ff
for dumbbell molecule, 116
for hard-sphere gas, 115
for Lennard-Jones gas, 115
for quantum gas, 119, 308
for rigid convex molecule, 117
Self-avoiding walk, 67, 207ff
Self-duality, 220
Short range order, 263
Solution, 150, 261
Specific heat, 15, 54, 76, 91, 94, 97, 180, 229, 243, 249, 255, 258, 259, 306, 392
electronic, 92, 94, 95, 100, 306
of Fermi liquids, 96
of Ising lattice, 229
of polyatomic gases, 16
of solids, 75ff, 242
Spectral broadening, 11
Spectral density, 200, 345
Spherical model, 246
Spin and statistics, 400
Spin correlation, 230ff
Spin wave, 247ff
Spontaneous magnetization, 229, 231, 234
Static form factor, 398
Statistical free energy, 35
Stirling's formula, 50, 105
Stokes friction, 159, 189
Stosszahlansatz, 4
Stress tensor, 157, 160, 370
Sum over states, 51
Superconductivity, 426ff
Superfluidity, 262, 331
Superposition approximation, 178
Surface internal energy, 164
Surface tension, 164
Susceptibility, 230, 232, 258, 260, 366

T

Terminal coupling, 222
Thermal conductivity, 15, 16, 99, 271, 340, 343, 345, 346, 352, 354
Thermal de Broglie wavelength, 62, 84, 283, 329
Thermodynamic limit, 137, 378
Thiele's semiinvariant, 124, 132, 145, 208
Third virial coefficient, 127, 128, 178
Thomas–Fermi screening constant, 325, 326, 395
Time arrow, 156
Time average, 25, 28, 29, 56
Tonks equation, 217
Toron, 291ff
Transformation group, 27
Transition probability, 42ff, 187, 206, 339, 349, 352, 356, 378
Transport phenomena, 13, 15, 99, 271, 338ff, 343ff, 360ff

U

Uehling–Uhlenbeck equation, 338ff
Umklapp process, 354

V

Van der Waals equation, 115
Velocity distribution function, 4ff, 280
Velocity gradient, 189, 370
Vertex, 414
Virial coefficient, 127
fourth, 178
second, 115, 147, 382
third, 128, 178
Virial equation of state, 163
Virial expansion, 126, 170
Viscosity, 14, 16, 17, 100, 158–160, 183, 341, 344, 350, 370
of normal fluid, 351
of quantum gas, 99, 343, 347, 370
of two-dimensional gas, 17

W

Watermelon approximation, 321, 324
Wave equation, 72
Weiss theory, 215, 265, 268
White spectrum, 202
Wick's time ordering operator, 406
Wiener–Khinchin theorem, 200, 379
Wigner distribution function, 280

X

X-ray scattering, 164, 167, 176

Z

Zero-point energy, 72, 75
Zero-point pressure, 94